REMARKS

Attached hereto is a marked up version of the changes made to the specification by this amendment. The attachment is captioned "**Version With Markings to Show Changes Made**."

Reconsideration and withdrawal of the rejections of this application and consideration and entry of this paper are respectfully requested in view of the herein remarks and accompanying information, which place the application in condition for allowance.

I. STATUS OF CLAIMS AND FORMAL MATTERS

Claims 1-25 are pending. Claims 1-25 are amended without prejudice, without admission, without surrender of subject matter, and without any intention of creating any estoppel as to equivalents.

No new matter is added. Support for the amended claims is found throughout the specification.

It is submitted that these claims are patentably distinct from the references cited by the Examiner, and that these claims are in full compliance with the requirements of 35 U.S.C. §112. The amendments of the claims herein are not made for the purpose of patentability within the meaning of 35 U.S.C. §§ 101, 102, 103 or 112; but rather the amendments are made simply for clarification and to round out the scope of protection to which Applicants are entitled.

II. RESPONSE TO RESTRICTION REQUIREMENT

The March 30, 2001 Office Action required an election under 35 U.S.C. § 121 and 37 CFR 1.141 from among:

Group I. SEQ ID NOs 1 and 7;

Group II. SEQ ID NOs 2 and 8;

Group III. SEQ ID NOs 3 and 9;

Group IV. SEQ ID NOs 4 and 10;

Group V. SEQ ID NOs 5 and 11; and

Group VI. SEQ ID NOs 6 and 12.

Pursuant to the telephone conversation of February 20, 2001 with Examiner Souaya, the election of Group I, SEQ ID NOs 1 and 7, with traverse, is affirmed. Reconsideration and withdrawal of the restriction requirement are respectfully requested in view of the remarks herewith.

As a traverse, it is noted that the MPEP lists two criteria for a proper restriction requirement. First, the inventions must be independent or distinct. MPEP § 803. Second, searching the additional inventions must constitute an undue burden on the examiner if restriction is not required. *Id.* The MPEP directs the examiner to search and examine an entire application "[i]f the search and examination of an entire application can be made without serious burden, ...even though it includes claims to distinct or independent inventions." *Id.*

The amino acid sequences (SEQ ID NOs 1-6) and the nucleic acids encoding them (SEQ ID NOs 7-12) are drawn functionally to the same enzyme, a glucan lyase. Furthermore, the invention is drawn to a process of producing an anti-oxidant *in situ*, in which process, preferably, an α-1,4-glucan lyase is used. The sequences 1-6 are given as six possible α-1,4-glucan lyases, as mere examples. Thus, each sequence does not constitute a separate invention, as the invention is the process of producing an anti-oxidant *in situ* from a glucan substrate.

In addition, it is respectfully submitted that a search of all of the related nucleotide and amino acid sequences presented in this application also does not pose a serious burden to the examiner. For example, as shown by the enclosed comparison, the peptides represented by SEQ ID NOs 2, 5 and 6 and the corresponding nucleic acids of SEQ ID NOs 8, 11 and 12 share greater than 74% sequence homology to SEQ ID NOs 1 and 7 respectively, underscoring the interrelated nature of these molecules. It would be unduly burdensome on the Applicants to have to file divisional applications to each amino acid sequence and corresponding nucleotide sequence, particularly as a "broad" process claim is sought.

Therefore, Groups I, II, III, IV, V and VI can be searched and examined in this application as there is no undue or serious burden in searching and examining these claims together. Thus, the restriction requirement is improper; and, it is respectfully requested that the restriction requirement be reconsidered and withdrawn.

III. <u>THE REJECTION UNDER §101 IS OVERCOME</u>

Claims 19-25 stand rejected under 35 U.S.C. §101 as allegedly not reciting any steps involved in the process. The rejection is respectfully traversed. The amendment herein has placed these claims in condition for allowance, rendering the rejection moot. Simply, the "use" recitation is no longer in the claims. Consequently, reconsideration and withdrawal of the rejection is respectfully requested.

IV. <u>THE REJECTION UNDER §112, FIRST PARAGRAPH, IS OVERCOME</u>

Claims 1-18 stand rejected under 35 U.S.C. §112, first paragraph, as allegedly containing non-enabling subject matter. The Office Action contends that the application fails to teach the method of the claimed invention with regard to any recombinant enzyme expressed in any plant to produce any anti-oxidant from a glucan substrate. Applicants disagree. The instant invention is clearly enabled because a skilled artisan would readily understand how to make and use the invention. Further, possession of the invention at the time of filing did exist.

According to the Court of Appeals for the Federal Circuit in the case of *In re Wands*, 8 U.S.P.Q. 2d 1400 (Fed. Cir. 1988),

> Enablement is not precluded by the necessity for some experimentation such as routine screening. However, experimentation needed to practice the invention must not be undue experimentation. 'The key word is undue, not experimentation.' The determination of what constitutes undue experimentation in a given case requires the application of standard of reasonableness, having due regard for the nature of the invention and the state of the art. The test is not merely quantitative, since a considerable amount of experimentation is permissible, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed ... [Citations omitted].

Id. at 1404.

Against this background, determining whether undue experimentation is required to practice a claimed invention turns on weighing many factors summarized in *In re Wands*, 858 F.2d 731, 8 USPQ2d 1400 (Fed. Cir. 1988) For example, (1) the quantity of experimentation necessary; (2) the amount of direction or guidance presented; (3) the presence or absence of working examples of the invention; (4) the nature of the invention; (5) the state of the prior art; (6) the relative skill of those in the art; (7) the predictability or unpredictability of the art; and (8) the breadth of the claims.

Never mentioned enablement

Thus, the assertion in the Office Action that the instant invention does not provide enablement for expressing any recombinant enzyme in any plant to produce any anti-oxidant from a glucan substrate is misplaced because undue experimentation would not exist. Applying *Wands* to the instant facts, it is clear that enablement exists, to wit, *inter alia*, that the quantity of experimentation necessary is low; the amount of direction or guidance presented is high;

working examples are clearly present; the relative skill of those in the art is high; and the predictability of the art is also high.

A skilled artisan would readily understand that a variety of recombinant enzymes which act on a glucan substrate, such as starch, to produce an anti-oxidant, such as anhydrofructose, are known in the art. For example, the enzyme amyloglucoside can act on a glucan substrate to produce glucose, which is then further converted to the anti-oxidant ascorbic acid, for example by glucose oxidase. Thus, it would be unduly restrictive to limit the invention to either a specific enzyme and/or to a specific anti-oxidant.

from any source. not fully described

In addition, several plants, namely grape, potatoes, maize and guar, are, in fact, exemplified (see pages 23-31 of the published PCT specification). It is hereby asserted that the present invention does disclose the invention in a manner which is clear and concise and which would enable a skilled person to carry out the invention.

Thus, it is respectfully submitted that the assertion in the Office Action that the claims are not enabled is obviated. Consequently, the Section 112, first paragraph, rejection should be reconsidered and withdrawn; and, such relief is respectfully requested.

V. THE REJECTION UNDER §112, SECOND PARAGRAPH, IS OVERCOME

Claims 1 and 17-25 stand rejected under 35 U.S.C. §112, second paragraph, as allegedly being indefinite. The amendments submitted herewith place the claims in condition for allowance and have rendered the Section 112, second paragraph, rejection moot. Simply, the bases for the rejection no longer appear in the claims. Consequently, reconsideration and withdrawal of the rejection is believed to be in order and such action is respectfully requested.

moot

VI. THE REJECTION UNDER §103 IS OVERCOME

Claims 1-18 stand rejected under 35 U.S.C. 103(a) as allegedly being unpatentable over Yu(a) or Yu(b) in view of Poulsen and Ishida et al. and Perl et al. The cited documents fail to disclose, suggest, or motivate a skilled artisan to practice, the presently claimed invention. The rejection is respectfully traversed.

The present invention relates to advantageous procedures based on the recombinant expression of glucan-metabolising enzymes in plants to produce anti-oxidants *in situ*.

One aspect of the present invention is based on a strategy of producing, in a medium, anti-oxidants which are metabolites of glucan by expressing within a plant, or part thereof, a

recombinant enzyme which acts on a glucan substrate present within the medium (including the plant or part thereof) to yield the anti-oxidant. This is in direct contrast to previously known procedures for producing glucan metabolite anti-oxidants which either rely on the addition of purified enzymes to a glucan-containing medium or use microorganisms expressing the enzymes.

In another aspect of the present invention, anti-oxidants are produced *in situ* in plants transformed with enzymes that convert glucans to anti-oxidant metabolites. This may be used, for example, to improve the stress tolerance of the plants or to improve the efficiency of transformation with foreign nucleic acid constructs.

Claim 1 and dependent claims involve a medium comprising a plant or part thereof. Thus, the amended claims are novel over the prior art disclosures. Although Yu(a) and Yu(b) may relate to the expression of α-1,4-glucan lyase in a host organism, such that it is capable of producing 1,5-D-anhydrofructose (AF), there is no mention of that organism being a plant. Rather, Yu(a) and (b) refer specifically to microorganisms such as bacteria, yeast and fungi. Indeed, Yu (a) and (b) are essentially concerned with the expression of α-1,4-glucan lyase and its subsequent use in the large-scale production of AF.

The skilled person concerned with obtaining α-1,4-glucan lyase in large quantities, for use in the large-scale production of AF and compounds such as the antibiotic microthecin, would not use plants for such a purpose; and, the transformation of microorganisms fails to teach or suggest transformation of plants or parts thereof. Further, Yu(a) and (b) are not concerned with improving the properties of plants or plant materials by *in situ* expression of α-1,4-glucan lyase. Thus, there is no incentive, teaching, or suggestion in Yu(a) and (b) to express α-1,4-glucan lyase in plants or parts thereof to produce antioxidants *in situ* from a glucan substrate.

Further, it is asserted by the Office Action that Poulsen, Ishida *et al.* and Perl *et al.* merely teach the transformation of potatoes, maize and grapes, respectively. While the transformation of plants was within the art, the Office Action has not shown a document which motivates the skilled person to even attempt to express an enzyme, for example α-1,4-glucan lyase, in a plant to produce *in situ* an anti-oxidant, for example AF, from a glucan substrate.

& widely known in the art

The state of the law in the Federal Circuit requires that the suggestion or motivation to practice the claimed invention must be present in the cited art, and not gleaned from the Applicant's specification. Further, picking and choosing portions from a plethora of disparate

references in a hindsight attempt to formulate an obviousness rejection is prohibited. In the instant case, the only suggestion or motivation to provide for the recombinant expression of glucan-metabolizing enzymes in plants to produce anti-oxidants *in situ* is found in the Applicant's disclosure. Thus, the cited documents fail to teach or suggest the instant invention. A *prima facie* case of obviousness has not been made and the rejections of the claims have been obviated.

No - well known in the art at the time of filing

In addition, the arguments in the Office Action with regard to obviousness, particularly with regard to the ease of plant transformation, apparently run counter to the Section 112 rejection; and, it is further respectfully asserted that the Examiner cannot have it both ways.

did not say → Not Not enabled.

It is submitted therefore that none of the cited documents, either alone or in combination, would lead the skilled person to express, in a plant, a recombinant enzyme that acts on a glucan substrate to produce an anti-oxidant. Thus, Claims 1-18 are novel and inventive (non-obvious) over the cited art. Accordingly, reconsideration and withdrawal of the Section 103 rejection is believed to be in order and such action is respectfully requested.

CONCLUSION

In view of the remarks and amendments herewith, the application is believed to be in condition for allowance. Favorable reconsideration of the application and prompt issuance of a Notice of Allowance are earnestly solicited. The undersigned looks forward to hearing favorably from the Examiner at an early date.

Respectfully submitted,

FROMMER LAWRENCE & HAUG LLP
Attorneys for Applicant

Thomas J. Kowalski

Thomas J. Kowalski
Reg. No. 32,147
Tel: (212) 588-0800

Version with Markings to Show Changes Made

IN THE CLAIMS

1. A process for producing an anti-oxidant in a medium [comprising a component] which medium comprises[is] a plant or part thereof, and which process[method] comprises expressing in the plant[component] or part thereof a recombinant enzyme which acts on a glucan substrate present in the medium and/or the component or part thereof, to yield said anti-oxidant.

2. The[A] process according to claim 1, wherein the glucan comprises α-1,4 links.

3. The[A] process according to claim 2, wherein the glucan is starch.

4. The[A] process according to claim 1, wherein the enzyme is a glucan lyase.

5. The[A] process according to claim 4, wherein the enzyme is an α-1,4-glucan lyase.

6. The[A] process according to claim 5, wherein the enzyme is any one of the sequences shown as SEQ ID NOs[Nos] 1-6, or a variant, homologue or fragment thereof.

7. The[A] process according to claim 6, wherein the enzyme is any one of the sequences shown as SEQ ID NOs[Nos] 1-6.

8. The[A] process according to claim 4, wherein the enzyme is encoded by a nucleotide sequence comprising any one of the sequences shown as SEQ ID NOs[Nos] 7-12, or a variant, homolog[ue] or fragment thereof.

9. The[A] process according to claim 8, wherein the enzyme is encoded by a nucleotide sequence having any one of the sequences shown as SEQ ID NOs[Nos] 7-12.

10. The[A] process according to claim 1, wherein the anti-oxidant is produced in the plant component, or part thereof, and is then released into the medium.

11. The[A] process according to claim 1, wherein the anti-oxidant is anhydrofructose.

12. <u>The</u>[A] process according to claim 11, wherein the anti-oxidant is1,5-D-anhydrofrutose.

13. <u>The</u>[A] process according to claim 1, wherein the medium, is, or is used in the preparation of, a foodstuff.

14. <u>The</u>[A] process according to claim 13, wherein the foodstuff is a beverage.

15. <u>The</u>[A] process according to claim 14, wherein the beverage is an alcoholic beverage.

16. <u>The</u>[A] process according to claim 14, wherein the beverage is a wine.

17. <u>The</u>[A] process according to claim <u>1</u>[13], wherein the <u>plant</u>[component] is all or part of a grape.

18. <u>The</u>[A] process according to claim 17, wherein the <u>plant</u>[component] is all or part of a grape.

19. [Use of anhydrofructose as]<u>Method of imparting into a medium</u> an anti-oxidant <u>comprising the steps of adding into the medium</u>[for a medium comprising] at least one component which is a plant or part thereof, wherein <u>the anti-oxidant is anhydrofructose, and wherein</u> the anhydrofructose is prepared *in situ* in the medium.

20. [Use of anhydrofructose as a means for]<u>Method of</u> imparting or improving stress tolerance in a plant, <u>comprising administering anhydrofructose,</u> wherein the anhydrofructose is prepared *in situ* in the plant.

21. [Use of anhydrofructose as a means for]<u>Method of</u> imparting or improving the transformation of a grape, <u>comprising administering anhydrofructose,</u> wherein the anhydrofructose is prepared *in situ* in the grape.

swability

22. [Use of glucan lyase as a means for]<u>Method of</u> imparting or <u>improving stress tolerance</u> in a plant, <u>comprising administering glucan lyase,</u> wherein the glucan lyase is prepared *in situ* in the plant.

23. [Use of glucan lyase as a means for]Method of imparting or improving the transformation of a grape, comprising administering glucan lyase, wherein the glucan lyase is prepared *in situ* in the grape.

24. [Use of a nucleotide sequence coding for a glucan lyase as a means for]Method of imparting or improving stress tolerance in a plant, comprising adminstering a nucleotide sequence coding for a glucan lyase, wherein the nucleotide sequence is expressed *in situ* in the plant.

25. [Use of a nucleotide sequence coding for a glucan lyase as a means for]Method of imparting or improving the transformation of a grape, comprising administering a nucleotide sequence coding for a glucan lyase, wherein the nucleotide sequence is expressed *in situ* in the grape.

P341 US

Comparison of glucan lyase aminoacid sequences of Sequence ID No. 1 against No.2-6

Part 1. Summary of the homolodiges of Squence ID No. 1 gainst No.2-6

Seq. ID. No.	1	2	3	4	5	6
Homology (%)	100	77.1	26.8	26.8	74.6	77.9

Part 2. Pairwise comparisons of Seq No. 1 against No. 2-6

1. Sequence Id No. 1 (SIN1) aligned with Sequence Id No. 2 (SIN2)

```
Sequence 1: SIN1, (1088 residues)
Sequence 2: SIN2, (1091 residues)

using the parameters:
Comparison matrix: BLOSUM62
Number of alignments computed: 20
Gap open penalty: 12
Gap extension penalty: 4

77.1% identity in 1096 residues overlap; Score: 4534.0; Gap frequency: 1.2%


SIN1,             1 MFSTLAFVAPSALGASTFVGAEV-RSNVRIHSAFPAVHTATRKTNRLNVSMTALSDKQTA
SIN2,             1 MYPTLTFVAPSALGARTFTCVGIFRSHILIHSVVPAVRLAVRKSNRLNVSMSALFDKPTA
                    *  ** ********* **      **   ***  ***  * ** ******* ** ** **

SIN1,            60 TAGSTDNPDGIDYKTYDYVGVWGFSPLSNTNWFAAGSSTPGGITDWTATMNVNFDRIDNP
SIN2,            61 VTGGKDNPDNINYTTYDYVPVWRFDPLSNTNWFAAGSSTPGDIDDWTATMNVNFDRIDNP
                      *  **** * * ***** ** * **************** * ****************

SIN1,           120 SITVQHPVQVQVTSYNNNSYRVRFNPDGPIRDVTRGPILKQQLDWIRTQELSEGCDPGMT
SIN2,           121 SFTLEKPVQVQVTSYKNNCFRVRFNPDGPIRDVDRGPILQQQLNWIRKQEQSKGFDPKMG
                    * *   ********* **  ************* ***** *** *** ** * * ** *

SIN1,           180 FTSEGFLTFETKDLSVIIYGNFKTRVTRKSDGKVIMENDEVGTASSGNKCRGLMFVDRLY
SIN2,           181 FTKEGFLKFETKDLNVIIYGNFKTRVTRKRDGKGIMENNEVPAGSLGNKCRGLMFVDRLY
                    ** **** ****** ************** *** **** **   * **************

SIN1,           240 GNAIASVNKNFRNDAVKQEGFYGAGEVNCKYQDT------YILERTGIAMTNYNYDNLNY
SIN2,           241 GTAIASVNENYRNDPDRKEGFYGAGEVNCEFWDSEQNRNKYILERTGIAMTNYNYDNYNY
                    * ****** * ***    ***********   *       ***************** **

SIN1,           294 NQWDLRPPHHDGALNPDYYIPMYYAAPWLIVNGCAGTS-EQYSYGWFMDNVSQSYMNTGD
SIN2,           301 NQSDLIAPGYPS--DPNFYIPMYFAAPWVVVKGCSGNSDEQYSYGWFMDNVSQTYMNTGG
                    ** **  *       *  ***** ****  * ** * * ************** *****

SIN1,           353 TTWNSGQEDLAYMGAQYGPFDQHFVYGAGGGMECVVTAFSLLQGKEFENQVLNKRSVMPP
SIN2,           359 TSWNCGEENLAYMGAQCGPFDQHFVYGDGDGLEDVVQAFSLLQGKEFENQVLNKRAVMPP
                    * ** * * ******* ********** * * * ** ****************** ****

SIN1,           413 KYVFGFFQGVFGTSSLLRAHMPAGENNISVEEIVEGYQNNNFPFEGLAVDVDMQDNLRVF
SIN2,           419 KYVFGYFQGVFGIASLLREQRPEGGNNISVQEIVEGYQSNNFPLEGLAVDVDMQQDLRVF
                    ***** ******  ****   * * ***** ******* **** **********  ****

SIN1,           473 TTKGEFWTANRVGTGGDPNNRSVFEWAHDKGLVCQTNITCFLRNDNEGQDYEVNQTLRER
SIN2,           479 TTKIEFWTANKVGTGGDSNNKSVFEWAHDKGLVCQTNVTCFLRNDNGGADYEVNQTLREK
                    *** ****** ****** ** **************** ******** * **********

SIN1,           533 QLYTKNDSLTGTDFGMTDDGPSDAYIGHLDYGGGVECDALFPDWGRPDVAEWWGNNYKKL
SIN2,           539 GLYTKNDSLTNTNFGTTNDGPSDAYIGHLDYGGGGNCDALFPDWGRPGVAEWWGDNYSKL
                     ********* * ** * ****************  *********** ****** ** **

SIN1,           593 FSIGLDFVWQDMTVPAMMPHKIGDDINVKPDGNWPNADDPSNGQYNWKTYHPQVLVTDMR
SIN2,           599 FKIGLDFVWQDMTVPAMMPHKVGDAVDTRSPYGWPNENDPSNGRYNWKSYHPQVLVTDMR
                    * ******************* **         ***  ***** **** ***********
```

```
SIN1,           653 YENHGREPMVTQRNIHAYTLCESTRKEGIVENADTLTKFRRSYIISRGGYIGNQHFGGMW
SIN2,           659 YENHGREPMFTQRNMHAYTLCESTRKEGIVANADTLTKFRRSYIISRGGYIGNQHFGGMW
                    ********* **** *************** *****************************

SIN1,           713 VGDNSTTSNYIQMMIANNINMNMSCLPLVGSDIGGFTSYDNENQRTPCTGDLMVRYVQAG
SIN2,           719 VGDNSSSQRYLQMMIANIVNMNMSCLPLVGSDIGGFTSYDGRNV---CPGDLMVRFVQAG
                    *****    * ******  *********************  *    * ****** ****

SIN1,           773 CLLPWFRNHYDRWIESKDHGKDYQELYMYPNEMDTLRKFVEFRYRWQEVLYTAMYQNAAF
SIN2,           776 CLLPWFRNHYGRLVEGKQEGKYYQELYMYKDEMATLRKFIEFRYRWQEVLYTAMYQNAAF
                    ********** *  * *  ** *******  ** ***** ********************

SIN1,           833 GKPIIKAASMYNNDSNVRRAQNDHFLLGGHDGYRILCAPVVWENSTERELYLPVLTQWYK
SIN2,           836 GKPIIKAASMYDNDRNVRGAQDDHFLLGGHDGYRILCAPVVWENTTSRDLYLPVLTKWYK
                    *********** ** *** ** ********************** * * ******* ***

SIN1,           893 FGPDFDTKPLEGAMNGGDRIYNYPVPQSESPIFVREGAILPTRYTLNGENKSLNTYTDED
SIN2,           896 FGPDYDTKRLDSALDGGQMIKNYSVPQSDSPIFVREGAILPTRYTLDGSNKSMNTYTDKD
                    **** *** *  *  **  * ** **** ***************** * *** ***** *

SIN1,           953 PLVFEVFPLGNNRADGMCYLDDGGVTTNAEDNGKFSVVKVAAEQDGGTETITFTNDCYEY
SIN2,           956 PLVFEVFPLGNNRADGMCYLDDGGITTDAEDHGKFSVINVEALRKGVTTTIKFAYDTYQY
                    ************************ ** *** *****  * *   * * ** *  * * *

SIN1,          1013 VFGGPFYVRVRGAQSPSNIHVSSGAGSQDMKVSSATSRAALFNDGENGDFWVDQETDSLW
SIN2,          1016 VFDGPFYVRIRNLTTASKINVSSGAGEEDMTPTSANSRAALFSDGGVGEYWADNDTSSLW
                    ** ****** *     * * ******  **   ** ****** **  *  * *  * ***

SIN1,          1073 LKLPNVVLPDAVITIT
SIN2,          1076 MKLPNLVLQDAVITIT
                     **** ** *******
```

^^

2. **Sequence Id No. 1** (SIN1) **aligned with Sequence Id No. 3** (SIN3)

Identity : 286 (26.83%)
Similarity: 166 (15.57%)
Number of gaps inserted in GL1FULL: 59
Number of gaps inserted in MCSY1: 67
GL1FULL (=SIN1). Total number of residues: 1088.
MCSY1. (=SIN3). Total number of residues: 1066.

Comparison matrix : Structure-genetic matrix.
Open gap cost : 2
Unit gap cost : 2

The character to show that two aligned residues are identical is '|'
The character to show that two aligned residues are similar is '.'
Amino acids said to be 'similar' are: A,S,T; D,E; N,Q; R,K; I,L,M,V; F,Y,W

```
GL1FULL  - M--FS-TLAFVAPSALGASTFVGAEVRSNVRIHSAFPAVHTATRKTNRLN -47
           | || ||       .  |  . .   .|      | |. |. .
MCSY1    - MAGFSDPLNFCKAEDYYS---VALDWKGPQKI---I-GVDTTPPKSTKFP -43

GL1FULL  - VSMTALSDKQTATAGSTDNPDGIDY-KTYDYVGVWG--FSP-LSNTNWFA -93
              . .      |     . . .     ||   . |     . .
MCSY1    - KNWHGVNLRF---DDGT--LGVVQFIRPC----VWRVRYDPGFKTSDEYG -84

GL1FULL  - AGSSTPGGITDWTATMNVNFDRIDN-PSITVQHPVQVQVTSYNNNS-YRV -141
             .   .|. .|.     ..|   .|          |         .
MCSY1    - DENTRTI-VQDYMSTLS---NKLDTYRGLTWE-------TKCEDSGDF-F -122

GL1FULL  - RFNPDGPIRDVTRGPILKQQL-DWIRTQELSEGCDPGMTFTSEGFLTFET -190
           |     . |.     .. . |.| |         |        |
MCSY1    - TFS--SKVTAVEKSERTRNKVGDGLRIH-LWKS----P-FRIQVVRTLTP -164

GL1FULL  - -KDLSVIIYGNFKTRVTRKSDGKVIMENDEVGTASSGNKCRGLMFVDRLY -239
```

```
MCSY1     - LKDPYPI--PNVAAAEARVSD-KV------VWQTSP--K----TF--R-- -195

GL1FULL   - GNAIASVNKNFRN---DAVK--QEGFYGAGEVNCKYQDTYILERTGIAMT -284
MCSY1     - -KNLHPQHKMLKDTVLDIVKPGHGEYVGWGEMGG-IQ--FMKEPT-F-MN -239

GL1FULL   - NYNYDNLNYNQWDLRPPHHDGALNP-DYYIPMYYAAPWLI-VNGCAGTSE -332
MCSY1     - YFNFDNMQYQQ--VYA---QGALDSRE---PLYHSDPFYLDVNS---NPE -278

GL1FULL   - QYSYGWFMDNVSQSYMNT-GDTTWNSGQEDLAY--MGAQYGPFDQHFVYG -379
MCSY1     - H-------KNITATFIDNYSQIAIDFGKTNSGYIKLGTRYGGID-CYGIS -320

GL1FULL   - AGGGMECVVTAFSLLQGKEFENQVLNKRSVMPPKYVFGFFQGVFGTSSLL -429
MCSY1     - ADTVPE-IVRLYTGLVGR---SKL--K-----PRYILGAHQACYGYQQ-- -357

GL1FULL   - RAHMPAGENNI-SVEEIVEGYQNNNFPFEGLAVDVDMQDNLRVFTTKGEF -478
MCSY1     - -------ESDLYSV---VQQYRDCKFPLDGIHVDVDVQDGFRTFTTN--- -394

GL1FULL   - WTANRVGTGGDPNNRSVFE-WAHDKGLVCQTNITCFLRNDN-EGQDY--- -523
MCSY1     - --PHTF-----PNPKEMFTNL-RNNGIKCSTNITPVISINNREG-GYSTL -435

GL1FULL   - -E-VNQT--LRERQLYTKNDSLTGTD-----FG--------MTD-DGPSD -555
MCSY1     - LEGVDKKYFIMD-DRYTEGTSGNAKDVRYMYYGGGNKVEVDPNDVNGRPD -484

GL1FULL   - -----AYIGHLD---YG--GGV------ECDALFPDWGRPDVAEWWGNNY -589
MCSY1     - FKDNYDFPANFNSKQYPYHGGVSYGYGNGSAGFYPDLNRKEVRIWWGMQY -534

GL1FULL   - KKLFSIGLDFVWQDMTVPAMMPHK-IGDDINVKPDGNWPN---ADDPSNG -635
MCSY1     - KYLFDMGLEFVWQDMT--TPAIHTSYGD-----MKG-LPTRLLV--TSDS -574

GL1FULL   - QYNWKTYHPQVLVTDMRYENHGREPMVTQRNIHAYTLCESTRKEGI--VE -683
MCSY1     - ------------VTNA---S-EKKLAIETWALYSYNLHKATWH-GLSRLE -607

GL1FULL   - NADTLTKFRRSYIISRGGYIGNQHFGGMWVGDNSTTSNYIQMMIANNINM -733
MCSY1     - S----RKNKRNFILGRGSYAGAYRFAGLWTGDNASNWEFWKISVSQVLSL -653

GL1FULL   - NMSCLPLVGSDIGGFTSYDNEN--QRTPCTGDLMVRYVQAGCLLPWFRNH -781
MCSY1     - GLNGVCIAGSDTGGFEPYRDANGVEEKYCSPELLIRWYTGSFLLPWLRNH -703

GL1FULL   - YDRWIESKDHGKDYQELYMYPN--E--MD---------TL----RKFVEF -814
MCSY1     - YVK----KDR-KWFQEPYSYPKHLETHPELADQAWLYKSVLEICRYYVEL -748

GL1FULL   - RYRWQEVLYTAMYQNAAFGKPIIKAASMYNN-DSNVRRAQNDHFL----L -859
MCSY1     - RYSLIQLLYDCMFQNVVDGMPITRSMLLTDTEDTTFF-NESQKFLDNQYM -797
```

```
GL1FULL   - GGHDGYRILCAPVWWENS-TE-----RELYLPVLTQWY--KFGP-DFDT- -899
            | |  || ||..  |  |     |..|||.  ||     | | |
MCSY1     - AGDD---ILVAPIL--HSRKEIPGENRDVYLPLYHTWYPSNLRPWD-DQG -841

GL1FULL   - ----KPLEGAM--NGGDRIY---NYPVPQSESPIFVREGAILPTRYTLNG -940
                |.||    |   ||     | .  |  |..||||||.|      .
MCSY1     - VALGNPVEGGSVINYTARIVAPEDYNLFHSWPVYVREGAIIPQ---I-- -886

GL1FULL   - ENKSLNTYTDEDPLVFEVFPLGNNRADGMCYLDDGGVTTNA--EDNGKFS -988
            | .            . | ..| |  .     ||||| |.  .  ||
MCSY1     - EVRQWTGQGGANRIKFNIYP-GKDKEYCT-YLDDG-VSRDSAPED---LP -930

GL1FULL   - WKVAAEQD--GGTETIT--F-TNDCYEYVFG-GP----FYVRVRGAQSP -1028
             | . ||   |.| |.          | . |  |     . .|  .|.
MCSY1     - QYKETHEQSKVEGAE-IAKQIGKKTGYN-ISGTDPEAKGYHRKVAVTQTS -978

GL1FULL   - SNIHVSSGAGSQDMKVSS-ATSRAA-------L-----FNDG-------- -1057
                 . . . . |       |.        |    |
MCSY1     - KD---KTRTVTIEPKHNGYDPSKEVGDYYTIILWYAPGFDGSIVDVSKTT -1025

GL1FULL   - ENGDFWVDQE----TD--SLWLKLPNVVL--PDAV-ITIT    -1088
            | . |.      .|  .. . .   |.    .| || |
MCSY1     - VNVEGGVEHQVYKNSDLHTVVIDV-KEVIGTTKSVKITCTAA -1066
```

^^

3. Sequence Id No. 1 (SIN1) aligned with Sequence Id No. 4 (SIN4)

Identity : 287 (26.82%)
Similarity: 160 (14.95%)
Number of gaps inserted in GL1FULL: 64
Number of gaps inserted in MVSY1: 63
* ALIGNMENT OF TWO PROTEIN SEQUENCES. *
The two sequences to be aligned are:

GL1FULL (=SIN1). Total number of residues: 1088.
MVSY1 (=SIN4). Total number of residues: 1070.

Comparison matrix : Structure-genetic matrix.
Open gap cost : 2
Unit gap cost : 2

The character to show that two aligned residues are identical is '|'
The character to show that two aligned residues are similar is '.'
Amino acids said to be 'similar' are: A,S,T; D,E; N,Q; R,K; I,L,M,V; F,Y,W

```
GL1FULL   - MFSTLAFVAPSALGAS-TFVGAEVRSNVRIHSAFPAVH--TATRKTNRLN -47
            |         . |    | ||         . |    .   | |
MVSY1     - M---------AGLSDPLNFCKAE---DY-----YAAAKGWSGPQKIIRYD -33

GL1FULL   - VSMTALSDKQTATAGSTDN-P--DGIDYKTYDYVG--VWG--FSP-LSNT -89
             .      |    . . | |  ||        .|   ||   . | .  .
MVSY1     - QTP-PQGTKD-PKSWHAVNLPFDDGTMCVV-QFVRPCVWRVRYDPSVKTS -80

GL1FULL   - NWFAAGSSTPGGITDWTATMNVNFDRIDNPSITVQHPVQVQVTSYNNNS- -138
              .    .    . |. .|.  | |       .|         |..
MVSY1     - DEYGDENTRTI-VQDYMTTLVGNLDIF--RGLTW-------VSTLEDSGE -120
```

```
GL1FULL  - YRVRFNPDGPIRDVTRGPILKQQLDWIRTQELSEGCDPGMTFTSEGFL-- -186
           |  | .  ||        . ||     ||.    .|
MVSY1    - Y-YTFKSEVTAVDET---------E--RTR--NKVGD-GLKI----YLWK -151

GL1FULL  - -TFETKDLSVIIYGNFKTRVTRKSDGKVIMENDEVGTASSGNKCRGLMFV -235
            |  . ..    |.  |  | | | |..  |
MVSY1    - NPFRIQWRLL------TPLV---DPFPIP-N--VANATA----R---VA -182

GL1FULL  - DRLYGNAIA-SVNKNFR--NDAVKQEGF--YGAGEVNCKYQDTYIL--ER -278
           |.. ..  . ||      .|.        |         |.  |
MVSY1    - DKVWQTSPKTFRKNLHPQHKMLKDTVLDIIKPG------HGEYVGWGEM -226

GL1FULL  - TGIA----MT--NY-NYDNLNYNQWDLRPPHHDGALNP-DYYIPMYYAAP -320
           ||     |  || |.||..|.|  .    . |||   .  |.| . |
MVSY1    - GGIEFMKEPTFMNYFNFDNMQYQQ--VYA---QGALDSRE---PLYHSDP -268

GL1FULL  - WLI-VNGCAGTSEQYSYGWFMDNVSQSYMNT-GDTTWNSGQEDLAY--MG -366
           . . ||     |        |.. ...       .   |    |  .|
MVSY1    - FYLDVNS---NPEH-------KNITATFIDNYSQIAIDFGKTNSGYIKLG -308

GL1FULL  - AQYGPFDQHFVYGAGGGMECWTAFSLLQGKEFENQVLNKRSVMPPKYVF -416
           . || |  .   |    | .|  .. | |.     .  |    |.|.
MVSY1    - TRYGGID-CYGISADTVPE-IVRLYTGLVGR---SKL--K-----PRYIL -346

GL1FULL  - GFFQGVFGTSSLLRAHMPAGENNISVEEIVEGYQNNNFPFEGLAVDVDMQ -466
           |  | .|               |  .    .|  |   || .|| |||| |
MVSY1    - GAHQACYGYQQ---------ESDLHA--WQQYRDTKFPLDGLHVDVDFQ -385

GL1FULL  - DNLRVFTTKG-EFWTANRVGTGGDPNNRSVFEWAHDKGLVCQTNITCFL- -514
           || | |||   |   |      .     .     |. | ||||  .
MVSY1    - DNFRTFTTNPITF--PNPK----EMFT-NL----RNNGIKCSTNITPVIS -424

GL1FULL  - -RN--D-----NEGQD--Y--------EVN----QTLR------ERQL-Y -535
            |          ||| |  |        |     | .|         .
MVSY1    - IRDRPNGYSTLNEGYDKKYFIMDDRYTEGTSGDPQNVRYSFYGGGNPVEV -474

GL1FULL  - TKNDSLTGTDFG-----MTD----DGPSDAYIGHLDYGGGVECDALFPDW -576
           || . |||       |     | |     .   || |      .||
MVSY1    - NPNDVWARPDFGDNYDFPTNFNCKDYPYHGGVSY-GYGNGTP--GYYPDL -521

GL1FULL  - GRPDVAEWWGNNYKKLFSIGLDFVWQDMTVPAMMPHKIGDDINVKPDGNW -626
           | .|  ||| .|  || .||.|||||| ||            |        .
MVSY1    - NREEVRIWWGLQYEYLFNMGLEFVWQDMTTPA---------IH----SSY -558

GL1FULL  - PNADDPSNGQYNWKTYHPQVLVTDMRYENHG-REPMVTQRNIHAYTLCES -675
              |   |          .|||      |   .    .   . .| | .
MVSY1    - ---GDM-KG------LPTRLLVTADSVTNASEKKLAIESWALYSYNLHKA -598

GL1FULL  - TRKE-GIVENADTLTKFRRSYIISRGGYIGNQHFGGMWVGDNSTTSNYIQ -724
           |    | .|       | .| .|. || | |   | |.| |||..|  .
MVSY1    - TFHGLGRLES----RKNKRNFILGRGSYAGAYRFAGLWTGDNASTWEFWK -644

GL1FULL  - MMIANNINMNMSCLPLVGSDIGGFTSYDNE-NQRTPCTGDLMVRYVQAGC -773
           . ... . . .  . . ||| |||     |      |. .|..|.
MVSY1    - ISVSQVLSLGLNGVCIAGSDTGGFEPARTEIGEEKYCSPELLIRWYTGSF -694

GL1FULL  - LLPWFRNHYDRWIESKDHGKDYQELYMYPNEMDT---L-------R---- -809
           |||| ||||  .    ||  |  .|| | ||  ..|   |         .
```

```
MVSY1     - LLPWLRNHYVK----KDR-KWFQEPYAYPKHLETHPELADQAWLYKSVLE -739

GL1FULL   - --KF-VEFRYRWQEVLYTAMYQNAA----FGK-PIIKAA--SMYNNDSNV -849
            .. || ||   .|| |.||          .  ..  .  . . |.|.
MVSY1     - ICRYWVELRYSLIQLLYDCMFQNVVDGMPLARSMLLTDTEDTTFFNESQ- -788

GL1FULL   - RRAQNDHFLLGGHDGYRILCAPVVWENS-TE-----RELYLPVLTQWY-- -891
             .     .. | |  || ||..  | |    |..|||.  ||
MVSY1     - -KFLDNQYM-AGDD---ILVAPIL--HSRNEVPGENRDVYLPLFHTWYPS -831

GL1FULL   - KFGP-DFDT-----KPLEGAM--NGGDRIY---NYPVPQSESPIFVREGA -930
              | | |     |.||   |   ||    | .    |...||||
MVSY1     - NLRPWD-DQGVALGNPVEGGSVINYTARIVAPEDYNLFHNVVPVYIREGA -880

GL1FULL   - ILPT---RYTLNGENKSLNTYTDEDPLVFEVFPLGNNRADGMCYLDDGGV -977
           |.|   |   ||           |. | ..| |  .    ||||| |
MVSY1     - IIPQIQVR-QWIGE--G-----GPNPIKFNIYP-GKDKEYVT-YLDDG-V -919

GL1FULL   - TTNAEDNGKFSVVKVAAEQD--GGTET---ITF-----TNDCYEYVFGGP -1017
            .  .        . | ||   | .   ..    |||     |
MVSY1     - SRDSAPDD-LPQYREAYEQAKVEGKDVQKQLAVIQGNKTND-----FSAS -963

GL1FULL   - FYVR-VRGAQSPSNI-HVS---SGAGSQDMKVSS-ATSRAA-------LF -1054
              .  .|     |   |   . . . . |     |.        |.
MVSY1     - GIDKEAKGYHRKVSIKQESKDKTRTVTIEPKHNGYDPSKEVGNYYTIILW -1013

GL1FULL   - NDGE-NGDFWVD--QETDSL--WLK--------LPNVVLPDA--VI-T-- -1086
                |   || | |  .   .          |  ||.    || |
MVSY1     - YAPGFDGSI-VDVSQATVNIEGGVECEIFKNTGLHTVVV-NVKEVIGTTK -1061

GL1FULL   - ---IT     -1088
               ||
MVSY1     - SVKITCTTA -1070
```

^^^

4. Sequence Id No. 1 (SIN1) aligned with Sequence Id No. 5 (SIN5)

```
Sequence 1: SIN1, (1088 residues)
Sequence 2: SIN5, (1092 residues)

using the parameters:
Comparison matrix: BLOSUM62
Number of alignments computed: 20
Gap open penalty: 12
Gap extension penalty: 4
```

74.6% identity in 1096 residues overlap; Score: 4455.0; Gap frequency: 1.3%

```
SIN1,            1 MFSTLAFVAPSALGASTFVGAEVRSNVRIHSAFPAVHTATRKTNRLNVSMTALSDKQTAT
SIN5,            1 MFPTLTFIAPSALAASTFVGADIRSGIRIQSALPAVRNAVRRSKHYNVSMTALSDKQTAI
                   ** ** * ***** *******  **  ** ** ***  * *     ************* 

SIN1,           61 AGSTDNPDGIDYKTYDYVGVWGFSPLSNTNWFAAGSSTPGGITDWTATMNVNFDRIDNPS
SIN5,           61 SIGPDNPDGINYQNYDYIPVAGFTPLSNTNWYAAGSSTPGGITDWTATMNVKFDRIDNPS
                       ****** *  ***  * ** ******* ******************* ********

SIN1,          121 ITVQHPVQVQVTSYNNNSYRVRFNPDGPIRDVTRGPILKQQLDWIRTQELSEGCDPGMTF
SIN5,          121 YSNNHPVQIQVTSYNNNSFRIRFNPDGPIRDVSRGPILKQQLTWIRNQELAQGCNPNMSF
                       **** ********* * *********** ********* *** ***  ** * * *

SIN1,          181 TSEGFLTFETKDLSVIIYGNFKTRVTRKSDGKVIMENDEVGTASSGNKCRGLMFVDRLYG
SIN5,          181 SPEGFLSFETKDLNVIIYGNCKMRVTKK-DGYLVMENDECNSQSDGNKCRGLMYVDRLYG
                     **** ****** ****** * *** * **   *****    * ******** ******

SIN1,          241 NAIASVNKNFRNDAVKQEGFYGAGEVNCKYQD-----TYILERTGIAMTNYNYDNLNYNQ
SIN5,          240 NAIASVQTNFHKDTSRNEKFYGAGEVNCRYEEQGKAPTYVLERSGLAMTNYNYDNLNYNQ
                   ******  **  *    * ********* *       ** *** * **************
```

```
SIN1,           296 WDLRPPHHDGALNPDYYIPMYYAAPWLIVNGCAGTSEQYSYGWFMDNVSQSYMNTGDTTW
SIN5,           300 PDVVPPGYPD--HPNYYIPMYYAAPWLVVQGCAGTSKQYSYGWFMDNVSQSYMNTGDTAW
                     *  **       * ************ * ****** ********************* *

SIN1,           356 NSGQEDLAYMGAQYGPFDQHFVYGAGGGMECVVTAFSLLQGKEFENQVLNKRSVMPPKYV
SIN5,           358 NCGQENLAYMGAQYGPFDQHFVYGDGDGLEDVVKAFSFLQGKEFEDKKLNKRSVMPPKYV
                    * *** ****************** * * * ** *** *******   ************

SIN1,           416 FGFFQGVFGTSSLLRAHMPAGENNISVEEIVEGYQNNNFPFEGLAVDVDMQDNLRVFTTK
SIN5,           418 FGFFQGVFGALSLLKQNLPAGENNISVQEIVEGYQDNDYPFEGLAVDVDMQDDLRVFTTK
                    *********  ***    ********* ******* *  ************* *******

SIN1,           476 GEFWTANRVGTGGDPNNRSVFEWAHDKGLVCQTNITCFLRNDNEGQDYEVNQTLRERQLY
SIN5,           478 PEYWSANMVGEGGDPNNRSVFEWAHDRGLVCQTNVTCFLRNDNSGKPYEVNQTLREKQLY
                     * * ** ** *************** ******* ******** *  ********* ***

SIN1,           536 TKNDSLTGTDFGMTDDGPSDAYIGHLDYGGGVECDALFPDWGRPDVAEWWGNNYKKLFSI
SIN5,           538 TKNDSLNNTDFGTTSDGPGDAYIGHLDYGGGVECDAIFPDWGRPDVAQWWGENYKKLFSI
                    ******  **** * *** ***************** ********** *** ********

SIN1,           596 GLDFVWQDMTVPAMMPHKIGDDINVKPDGN---WPNADDPSNGQYNWKTYHPQVLVTDMR
SIN5,           598 GLDFVWQDMTVPAMMPHRLGDAVNKNSGSSAPGWPNENDPSNGRYNWKSYHPQVLVTDMR
                    *****************  **  *         ***  ***** **** ***********

SIN1,           653 Y-ENHGREPMVTQRNIHAYTLCESTRKEGIVENADTLTKFRRSYIISRGGYIGNQHFGGM
SIN5,           658 YGAEYGREPMVSQRNIHAYTLCESTRREGIVGNADSLTKFRRSYIISRGGYIGNQHFGGM
                    *    ****** ************** **** *** ************************

SIN1,           712 WVGDNSTTSNYIQMMIANNINMNMSCLPLVGSDIGGFTSYDNENQRTPCTGDLMVRYVQA
SIN5,           718 WVGDNSATESYLQMMLANIINMNMSCLPLVGSDIGGFTQYNDAGDPTP--EDLMVRFVQA
                    ****** *  * *** ** ******************* *      **   ***** ***

SIN1,           772 GCLLPWFRNHYDRWIESKDHGKDYQELYMYPNEMDTLRKFVEFRYRWQEVLYTAMYQNAA
SIN5,           776 GCLLPWFRNHYDRWIESKKHGKKYQELYMYPGQKDTLKKFVEFRYRWQEVLYTAMYQNAT
                    ****************** *** ********   *** *********************

SIN1,           832 FGKPIIKAASMYNNDSNVRRAQNDHFLLGGHDGYRILCAPVVWENSTERELYLPVLTQWY
SIN5,           836 TGEPIIKAAPMYNNDVNVYKSQNDHFLLGGHDGYRILCAPVVRENATSREVYLPVYSKWF
                     * ****** ***** **   ********************* ** * ** ****   *

SIN1,           892 KFGPDFDTKPLEGAMNGGDRIYNYPVPQSESPIFVREGAILPTRYTLNGENKSLNTYTDE
SIN5,           896 KFGPDFDTKPLENEIQGGQTLYNYAAPLNDSPIFVREGTILPTRYTLDGVNKSINTYTDN
                    ************    **   ***  *   ******** ******** * *** *****

SIN1,           952 DPLVFEVFPLGNNRADGMCYLDDGGVTTNAEDNGKFSVVKVAAEQDGGTETITFTNDCYE
SIN5,           956 DPLVFELFPLENNQAHGLFYHDDGGVTTNAEDFGKYSVISVKAAQEGSQMSVKFDNEVYE
                    ****** *** ** * *  * *********** ** **  * * * *      * *  **

SIN1,          1012 YVFGGPFYVRVRGAQSPSNIHVSSGAGSQDMKVSSATSRAALFNDGENGDFWVDQETDSL
SIN5,          1016 HQWGASFYVRVRNMGAPSNINVSSQIGQQDMQQSSVSSRAQMFTSANDGEYWVDQSTNSL
                       *  ******    **** ***  * ***  **  ***  *     *  **** * **

SIN1,          1072 WLKLPNVVLPDAVITI
SIN5,          1076 WLKLPGAVIQDAAITV
                    *****  *  ** **
```

^^

5. Sequence Id No. 1 (SIN1) aligned with Sequence Id No. 6 (SIN6)

Sequence 1: SIN1, (1088 residues)
Sequence 2: SIN6, (570 residues)

using the parameters:
Comparison matrix: BLOSUM62
Number of alignments computed: 20
Gap open penalty: 12
Gap extension penalty: 4
77.9% identity in 578 residues overlap; Score: 2451.0; Gap frequency: 1.9%

```
SIN1,           283 MTNYNYDNLNYNQWDLRPPHHDGALNPDYYIPMYYAAPWLIVNGCAGTSEQYSYGWFMDN
SIN6,             1 MTNYNYDNLNYNQPDLIPPGHDS--DPDYYIPMYFAAPWVIAHGYRGTSDQYSYGWFLDN
                    ************* ** ** **    ******** **** *  *  *** ******* **

SIN1,           343 VSQSYMNTGDTTWNSGQEDLAYMGAQYGPFDQHFVYGAGGGMECVVTAFSLLQGKEFENQ
SIN6,            59 VSQSYTNTGDDAW-AGQKDLAYMGAQCGPFDQHFVYEAGDGLEDVVTAFSYLQGKEYENQ
                    ***** ****  *  ** ******** ********* ** * * ****** ***** ***

SIN1,           403 VLNKRSVMPPKYVFGFFQGVFGTSSLLRAHMPAGENNISVEEIVEGYQNNNFPFEGLAVD
```

```
SIN6,     118 GLNIRSAMPPKYVFGFFQGVFGATSLLRDNLPAGENNVSLEEIVEGYQNQNVPFEGLAVD
               ** ** *************** ****   ****** * ********* * ********

SIN1,     463 VDMQDNLRVFTTKGEFWTANRVGTGGDPNNRSVFEWAHDKGLVCQTNITCFLRNDNEGQD
SIN6,     178 VDMQDDLRVFTTRPAFWTANKVGEGGDPNNKSVFEWAHDRGLVCQTNVTCFLK--NEKNP
              ***** ******   ***** ** ****** ******** ******* ****   **

SIN1,     523 YEVNQTLRERQLYTKNDSLTGTDFGMTDDGPSDAYIGHLDYGGGVECDALFPDWGRPDVA
SIN6,     236 YEVNQSLREKQLYTKSDSLDNIDFGTTPDGPSDAYIGHLDYGGGVECDALFPDWGRPDVA
              ***** *** ***** ***    *** * ********************************

SIN1,     583 EWWGNNYKKLFSIGLDFVWQDMTVPAMMPHKIGDDINVKPDGN---WPNADDPSNGQYNW
SIN6,     296 QWWGDNYKKLFSIGLDFVWQDMTVPAMMPHRLGDPVGTNSGETAPGWPNDKDPSNGRYNW
               *** *************************  **            ***  ***** ***

SIN1,     640 KTYHPQVLVTDMRYENHGREPMVTQRNIHAYTLCESTRKEGIVENADTLTKFRRSYIISR
SIN6,     356 KSYHPQVLVTDMRYDDYGRDPIVTQRNLHAYTLCESTRREGIVGNADSLTKFRRSYIISR
              * ************    ** * ***** ********** **** *** ************

SIN1,     700 GGYIGNQHFGGMWVGDNSTTSNYIQMMIANNINMNMSCLPLVGSDIGGFTSYDNENQRTP
SIN6,     416 GGYIGNQHFGGMWVGDNSSTEDYLAMMVINVINMNMSGVPLVGSDIGGFTEHDKRN---P
              ****************** *  *  **  * ******  ***********  *  *   *

SIN1,     760 CTGDLMVRYVQAGCLLPWFRNHYDRWIESKDHGKDYQELYMYPNEMDTLRKFVEFRYRWQ
SIN6,     473 CTPDLMMRFVQAGCLLPWFRNHYDRWIESKKHGKNYQELYMYRDHLDALRSFVELRYRWQ
              ** *** * ********************* *** *******    * ** *** *****

SIN1,     820 EVLYTAMYQNAAFGKPIIKAASMYNNDSNVRRAQNDHF
SIN6,     533 EVLYTAMYQNALNGKPIIKTVSMYNNDMNVKDAQNDHF
              ***********  ******  ****** **  ******
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
```

Comparison of glucan lyase aminoacid sequences of Sequence ID No. 1 against No.2-6

Part 1. Summary of the homolodiges of Squence ID No. 1 gainst No.2-6

Seq. ID. No.	1	2	3	4	5	6
Homology (%)	100	77.1	26.8	26.8	74.6	77.9

Part 2. Pairwise comparisons of Seq No. 1 against No. 2-6

1. Sequence Id No. 1 (SIN1) aligned with Sequence Id No. 2 (SIN2)

```
Sequence 1: SIN1, (1088 residues)
Sequence 2: SIN2, (1091 residues)

using the parameters:
Comparison matrix: BLOSUM62
Number of alignments computed: 20
Gap open penalty: 12
Gap extension penalty: 4
```

77.1% identity in 1096 residues overlap; Score: 4534.0; Gap frequency: 1.2%

```
SIN1,         1 MFSTLAFVAPSALGASTFVGAEV-RSNVRIHSAFPAVHTATRKTNRLNVSMTALSDKQTA
SIN2,         1 MYPTLTFVAPSALGARTFTCVGIFRSHILIHSVVPAVRLAVRKSNRLNVSMSALFDKPTA
                *  ** ********* **      **   ***  ***  * ** ******* ** ** **

SIN1,        60 TAGSTDNPDGIDYKTYDYVGVWGFSPLSNTNWFAAGSSTPGGITDWTATMNVNFDRIDNP
SIN2,        61 VTGGKDNPDNINYTTYDYVPVWRFDPLSNTNWFAAGSSTPGDIDDWTATMNVNFDRIDNP
                  *  **** * * ***** ** * **************** * ****************

SIN1,       120 SITVQHPVQVQVTSYNNNSYRVRFNPDGPIRDVTRGPILKQQLDWIRTQELSEGCDPGMT
SIN2,       121 SFTLEKPVQVQVTSYKNNCFRVRFNPDGPIRDVDRGPILQQQLNWIRKQEQSKGFDPKMG
                * *   ********* **  ************* ***** *** *** ** * * ** *

SIN1,       180 FTSEGFLTFETKDLSVIIYGNFKTRVTRKSDGKVIMENDEVGTASSGNKCRGLMFVDRLY
SIN2,       181 FTKEGFLKFETKDLNVIIYGNFKTRVTRKRDGKGIMENNEVPAGSLGNKCRGLMFVDRLY
                ** **** ****** ************** *** **** **   * **************

SIN1,       240 GNAIASVNKNFRNDAVKQEGFYGAGEVNCKYQDT------YILERTGIAMTNYNYDNLNY
SIN2,       241 GTAIASVNENYRNDPDRKEGFYGAGEVNCEFWDSEQNRNKYILERTGIAMTNYNYDNYNY
                *  ****** * ***    ***********   *       ***************** **

SIN1,       294 NQWDLRPPHHDGALNPDYYIPMYYAAPWLIVNGCAGTS-EQYSYGWFMDNVSQSYMNTGD
SIN2,       301 NQSDLIAPGYPS--DPNFYIPMYFAAPWVVVKGCSGNSDEQYSYGWFMDNVSQTYMNTGG
                ** **  *       *  ***** ****  * ** * * ************** *****

SIN1,       353 TTWNSGQEDLAYMGAQYGPFDQHFVYGAGGGMECVVTAFSLLQGKEFENQVLNKRSVMPP
SIN2,       359 TSWNCGEENLAYMGAQCGPFDQHFVYGDGDGLEDVVQAFSLLQGKEFENQVLNKRAVMPP
                * ** * * ******* ********** * * * ** ****************** ****

SIN1,       413 KYVFGFFQGVFGTSSLLRAHMPAGENNISVEEIVEGYQNNNFPFEGLAVDVDMQDNLRVF
SIN2,       419 KYVFGYFQGVFGIASLLREQRPEGGNNISVQEIVEGYQSNNFPLEGLAVDVDMQQDLRVF
                ***** ******  ****   * * ***** ******* **** **********  ****

SIN1,       473 TTKGEFWTANRVGTGGDPNNRSVFEWAHDKGLVCQTNITCFLRNDNEGQDYEVNQTLRER
SIN2,       479 TTKIEFWTANKVGTGGDSNNKSVFEWAHDKGLVCQTNVTCFLRNDNGGADYEVNQTLREK
                *** ****** ****** ** **************** ******** * **********

SIN1,       533 QLYTKNDSLTGTDFGMTDDGPSDAYIGHLDYGGGVECDALFPDWGRPDVAEWWGNNYKKL
SIN2,       539 GLYTKNDSLTNTNFGTTNDGPSDAYIGHLDYGGGGNCDALFPDWGRPGVAEWWGDNYSKL
                 ********* * ** * ****************  *********** ****** ** **

SIN1,       593 FSIGLDFVWQDMTVPAMMPHKIGDDINVKPDGNWPNADDPSNGQYNWKTYHPQVLVTDMR
SIN2,       599 FKIGLDFVWQDMTVPAMMPHKVGDAVDTRSPYGWPNENDPSNGRYNWKSYHPQVLVTDMR
                * ******************* **         ***  ***** **** ***********
```

```
SIN1,       653 YENHGREPMVTQRNIHAYTLCESTRKEGIVENADTLTKFRRSYIISRGGYIGNQHFGGMW
SIN2,       659 YENHGREPMFTQRNMHAYTLCESTRKEGIVANADTLTKFRRSYIISRGGYIGNQHFGGMW
                ********* **** *************** *****************************

SIN1,       713 VGDNSTTSNYIQMMIANNINMNMSCLPLVGSDIGGFTSYDNENQRTPCTGDLMVRYVQAG
SIN2,       719 VGDNSSSQRYLQMMIANIVNMNMSCLPLVGSDIGGFTSYDGRNV---CPGDLMVRFVQAG
                *****    * ******  *********************  *    * ****** ****

SIN1,       773 CLLPWFRNHYDRWIESKDHGKDYQELYMYPNEMDTLRKFVEFRYRWQEVLYTAMYQNAAF
SIN2,       776 CLLPWFRNHYGRLVEGKQEGKYYQELYMYKDEMATLRKFIEFRYRWQEVLYTAMYQNAAF
                ********** *  * *  ** *******  ** ***** ********************

SIN1,       833 GKPIIKAASMYNNDSNVRRAQNDHFLLGGHDGYRILCAPVVWENSTERELYLPVLTQWYK
SIN2,       836 GKPIIKAASMYDNDRNVRGAQDDHFLLGGHDGYRILCAPVVWENTTSRDLYLPVLTKWYK
                *********** ** *** ** ********************** * * ******* ***

SIN1,       893 FGPDFDTKPLEGAMNGGDRIYNYPVPQSESPIFVREGAILPTRYTLNGENKSLNTYTDED
SIN2,       896 FGPDYDTKRLDSALDGGQMIKNYSVPQSDSPIFVREGAILPTRYTLDGSNKSMNTYTDKD
                **** *** *  *  **  * ** **** ***************** * *** ***** *

SIN1,       953 PLVFEVFPLGNNRADGMCYLDDGGVTTNAEDNGKFSVVKVAAEQDGGTETITFTNDCYEY
SIN2,       956 PLVFEVFPLGNNRADGMCYLDDGGITTDAEDHGKFSVINVEALRKGVTTTIKFAYDTYQY
                ************************ ** *** *****  * *   *  * ** *  *  * *

SIN1,      1013 VFGGPFYVRVRGAQSPSNIHVSSGAGSQDMKVSSATSRAALFNDGENGDFWVDQETDSLW
SIN2,      1016 VFDGPFYVRIRNLTTASKINVSSGAGEEDMTPTSANSRAALFSDGGVGEYWADNDTSSLW
                ** ****** *     * * ******  **   ** ****** **  *  * *  *  ***

SIN1,      1073 LKLPNVVLPDAVITIT
SIN2,      1076 MKLPNLVLQDAVITIT
                 **** ** *******
```

^^

2. **Sequence Id No. 1** (SIN1) **aligned with Sequence Id No. 3** (SIN3)

Identity : 286 (26.83%)

Similarity: 166 (15.57%)
Number of gaps inserted in GL1FULL: 59
Number of gaps inserted in MCSY1: 67
GL1FULL (=SIN1). Total number of residues: 1088.
MCSY1. (=SIN3). Total number of residues: 1066.

Comparison matrix : Structure-genetic matrix.
Open gap cost : 2
Unit gap cost : 2

The character to show that two aligned residues are identical is '|'
The character to show that two aligned residues are similar is '.'
Amino acids said to be 'similar' are: A,S,T; D,E; N,Q; R,K; I,L,M,V; F,Y,W

```
GL1FULL   - M--FS-TLAFVAPSALGASTFVGAEVRSNVRIHSAFPAVHTATRKTNRLN -47
            |  ||  | |       .   |  . .   .|      | |.  |. .
MCSY1     - MAGFSDPLNFCKAEDYYS---VALDWKGPQKI---I-GVDTTPPKSTKFP -43

GL1FULL   - VSMTALSDKQTATAGSTDNPDGIDY-KTYDYVGVWG--FSP-LSNTNWFA -93
                 .  .       |     . . .      ||   . |    .  .
MCSY1     - KNWHGVNLRF---DDGT--LGVVQFIRPC----VWRVRYDPGFKTSDEYG -84

GL1FULL   - AGSSTPGGITDWTATMNVNFDRIDN-PSITVQHPVQVQVTSYNNNS-YRV -141
               .    . |. .|.     ..|    .|         |       .
MCSY1     - DENTRTI-VQDYMSTLS---NKLDTYRGLTWE-------TKCEDSGDF-F -122

GL1FULL   - RFNPDGPIRDVTRGPILKQQL-DWIRTQELSEGCDPGMTFTSEGFLTFET -190
             |     .  | .    .. . | .|   |         |      |
MCSY1     - TFS--SKVTAVEKSERTRNKVGDGLRIH-LWKS----P-FRIQVVRTLTP -164

GL1FULL   - -KDLSVIIYGNFKTRVTRKSDGKVIMENDEVGTASSGNKCRGLMFVDRLY -239
```

```
MCSY1     - LKDPYPI--PNVAAAEARVSD-KV------VWQTSP--K----TF--R-- -195

GL1FULL   - GNAIASVNKNFRN---DAVK--QEGFYGAGEVNCKYQDTYILERTGIAMT -284
MCSY1     - -KNLHPQHKMLKDTVLDIVKPGHGEYVGWGEMGG-IQ--FMKEPT-F-MN -239

GL1FULL   - NYNYDNLNYNQWDLRPPHHDGALNP-DYYIPMYYAAPWLI-VNGCAGTSE -332
MCSY1     - YFNFDNMQYQQ--VYA---QGALDSRE---PLYHSDPFYLDVNS---NPE -278

GL1FULL   - QYSYGWFMDNVSQSYMNT-GDTTWNSGQEDLAY--MGAQYGPFDQHFVYG -379
MCSY1     - H-------KNITATFIDNYSQIAIDFGKTNSGYIKLGTRYGGID-CYGIS -320

GL1FULL   - AGGGMECVVTAFSLLQGKEFENQVLNKRSVMPPKYVFGFFQGVFGTSSLL -429
MCSY1     - ADTVPE-IVRLYTGLVGR---SKL--K-----PRYILGAHQACYGYQQ-- -357

GL1FULL   - RAHMPAGENNI-SVEEIVEGYQNNNFPFEGLAVDVDMQDNLRVFTTKGEF -478
MCSY1     - -------ESDLYSV---VQQYRDCKFPLDGIHVDVDVQDGFRTFTTN--- -394

GL1FULL   - WTANRVGTGGDPNNRSVFE-WAHDKGLVCQTNITCFLRNDN-EGQDY--- -523
MCSY1     - --PHTF-----PNPKEMFTNL-RNNGIKCSTNITPVISINNREG-GYSTL -435

GL1FULL   - -E-VNQT--LRERQLYTKNDSLTGTD-----FG--------MTD-DGPSD -555
MCSY1     - LEGVDKKYFIMD-DRYTEGTSGNAKDVRYMYYGGGNKVEVDPNDVNGRPD -484

GL1FULL   - -----AYIGHLD---YG--GGV------ECDALFPDWGRPDVAEWWGNNY -589
MCSY1     - FKDNYDFPANFNSKQYPYHGGVSYGYGNGSAGFYPDLNRKEVRIWWGMQY -534

GL1FULL   - KKLFSIGLDFVWQDMTVPAMMPHK-IGDDINVKPDGNWPN---ADDPSNG -635
MCSY1     - KYLFDMGLEFVWQDMT--TPAIHTSYGD-----MKG-LPTRLLV--TSDS -574

GL1FULL   - QYNWKTYHPQVLVTDMRYENHGREPMVTQRNIHAYTLCESTRKEGI--VE -683
MCSY1     - -------------VTNA---S-EKKLAIETWALYSYNLHKATWH-GLSRLE -607

GL1FULL   - NADTLTKFRRSYIISRGGYIGNQHFGGMWVGDNSTTSNYIQMMIANNINM -733
MCSY1     - S----RKNKRNFILGRGSYAGAYRFAGLWTGDNASNWEFWKISVSQVLSL -653

GL1FULL   - NMSCLPLVGSDIGGFTSYDNEN--QRTPCTGDLMVRYVQAGCLLPWFRNH -781
MCSY1     - GLNGVCIAGSDTGGFEPYRDANGVEEKYCSPELLIRWYTGSFLLPWLRNH -703

GL1FULL   - YDRWIESKDHGKDYQELYMYPN--E--MD----------TL----RKFVEF -814
MCSY1     - YVK----KDR-KWFQEPYSYPKHLETHPELADQAWLYKSVLEICRYYVEL -748

GL1FULL   - RYRWQEVLYTAMYQNAAFGKPIIKAASMYNN-DSNVRRAQNDHFL----L -859
MCSY1     - RYSLIQLLYDCMFQNVVDGMPITRSMLLTDTEDTTFF-NESQKFLDNQYM -797
```

```
GL1FULL  - GGHDGYRILCAPVVWENS-TE-----RELYLPVLTQWY--KFGP-DFDT- -899
           | |  || ||..  | |    |..|||.  ||    | | |
MCSY1    - AGDD---ILVAPIL--HSRKEIPGENRDVYLPLYHTWYPSNLRPWD-DQG -841

GL1FULL  - ----KPLEGAM--NGGDRIY---NYPVPQSESPIFVREGAILPTRYTLNG -940
               |.||   |  ||    | .  |  |..||||||.|    .
MCSY1    - VALGNPVEGGSVINYTARIVAPEDYNLFHSWPVYVREGAIIPQ---I-- -886

GL1FULL  - ENKSLNTYTDEDPLVFEVFPLGNNRADGMCYLDDGGVTTNA--EDNGKFS -988
           | .          . | ..| |  .    ||||| |.  .  ||
MCSY1    - EVRQWTGQGGANRIKFNIYP-GKDKEYCT-YLDDG-VSRDSAPED---LP -930

GL1FULL  - VVKVAAEQD--GGTETIT--F-TNDCYEYVFG-GP----FYVRVRGAQSP -1028
            | . ||   |.| |.          |  . |  |     .  .|  .|.
MCSY1    - QYKETHEQSKVEGAE-IAKQIGKKTGYN-ISGTDPEAKGYHRKVAVTQTS -978

GL1FULL  - SNIHVSSGAGSQDMKVSS-ATSRAA-------L-----FNDG--------- -1057
                  . . . . |       |.          |     |
MCSY1    - KD---KTRTVTIEPKHNGYDPSKEVGDYYTIILWYAPGFDGSIVDVSKTT -1025

GL1FULL  - ENGDFWVDQE----TD--SLWLKLPNVVL--PDAV-ITIT    -1088
           | .  |.      .|  .. . .   |.    .| || |
MCSY1    - VNVEGGVEHQVYKNSDLHTVVIDV-KEVIGTTKSVKITCTAA -1066
```

^^

3. Sequence Id No. 1 (SIN1) aligned with Sequence Id No. 4 (SIN4)

Identity : 287 (26.82%)
Similarity: 160 (14.95%)
Number of gaps inserted in GL1FULL: 64
Number of gaps inserted in MVSY1: 63
* ALIGNMENT OF TWO PROTEIN SEQUENCES. *
The two sequences to be aligned are:

GL1FULL (=SIN1). Total number of residues: 1088.
MVSY1 (=SIN4). Total number of residues: 1070.

Comparison matrix : Structure-genetic matrix.
Open gap cost : 2
Unit gap cost : 2

The character to show that two aligned residues are identical is '|'
The character to show that two aligned residues are similar is '.'
Amino acids said to be 'similar' are: A,S,T; D,E; N,Q; R,K; I,L,M,V; F,Y,W

```
GL1FULL  - MFSTLAFVAPSALGAS-TFVGAEVRSNVRIHSAFPAVH--TATRKTNRLN -47
           |          . |    |  ||           . |     .   |  |
MVSY1    - M----------AGLSDPLNFCKAE---DY-----YAAAKGWSGPQKIIRYD -33

GL1FULL  - VSMTALSDKQTATAGSTDN-P--DGIDYKTYDYVG--VWG--FSP-LSNT -89
            .      |    .  . | |  ||        .|   ||   . | .  .
MVSY1    - QTP-PQGTKD-PKSWHAVNLPFDDGTMCVV-QFVRPCVWRVRYDPSVKTS -80

GL1FULL  - NWFAAGSSTPGGITDWTATMNVNFDRIDNPSITVQHPVQVQVTSYNNNS- -138
            .    .     . |. .|.  | |        .|          |..
MVSY1    - DEYGDENTRTI-VQDYMTTLVGNLDIF--RGLTW-------VSTLEDSGE -120
```

```
GL1FULL  - YRVRFNPDGPIRDVTRGPILKQQLDWIRTQELSEGCDPGMTFTSEGFL-- -186
           |  |  .   ||          .  ||        | |.      .|
MVSY1    - Y-YTFKSEVTAVDET---------E--RTR--NKVGD-GLKI----YLWK -151

GL1FULL  - -TFETKDLSVIIYGNFKTRVTRKSDGKVIMENDEVGTASSGNKCRGLMFV -235
            |    . ..       | .    |    |   |   |   |  |..     |
MVSY1    - NPFRIQWRLL------TPLV---DPFPIP-N--VANATA----R---VA -182

GL1FULL  - DRLYGNAIA-SVNKNFR--NDAVKQEGF--YGAGEVNCKYQDTYIL--ER -278
           |..  ..   .  ||       .|          |           |.   |
MVSY1    - DKVWWQTSPKTFRKNLHPQHKMLKDTVLDIIKPG------HGEYVGWGEM -226

GL1FULL  - TGIA----MT--NY-NYDNLNYNQWDLRPPHHDGALNP-DYYIPMYYAAP -320
           ||      |  ||  |.||..|.|   .      |||    .    |.| . |
MVSY1    - GGIEFMKEPTFMNYFNFDNMQYQQ--VYA---QGALDSRE---PLYHSDP -268

GL1FULL  - WLI-VNGCAGTSEQYSYGWFMDNVSQSYMNT-GDTTWNSGGQEDLAY--MG -366
           . . ||       |          |.. ...          .    |      |   .|
MVSY1    - FYLDVNS---NPEH-------KNITATFIDNYSQIAIDFGKTNSGYIKLG -308

GL1FULL  - AQYGPFDQHFVYGAGGGMECVVTAFSLLQGKEFENQVLNKRSVMPPKYVF -416
           . ||  |   .   |      | .|  .. | |.       .  |       |.|.
MVSY1    - TRYGGID-CYGISADTVPE-IVRLYTGLVGR---SKL--K-----PRYIL -346

GL1FULL  - GFFQGVFGTSSLLRAHMPAGENNISVEEIVEGYQNNNFPFEGLAVDVDMQ -466
           |  |  .|                    |   .     .|  |    || .|| |||| |
MVSY1    - GAHQACYGYQQ---------ESDLHA--VVQQYRDTKFPLDGLHVDVDFQ -385

GL1FULL  - DNLRVFTTKG-EFWTANRVGTGGDPNNRSVFEWAHDKGLVCQTNITCFL- -514
           || | |||     |     |          .      .           |. | |||| .
MVSY1    - DNFRTFTTNPITF--PNPK----EMFT-NL----RNNGIKCSTNITPVIS -424

GL1FULL  - -RN--D-----NEGQD--Y--------EVN----QTLR------ERQL-Y -535
           |              ||| |   |            |        | .|           .
MVSY1    - IRDRPNGYSTLNEGYDKKYFIMDDRYTEGTSGDPQNVRYSFYGGGNPVEV -474

GL1FULL  - TKNDSLTGTDFG-----MTD----DGPSDAYIGHLDYGGGVECDALFPDW -576
            ||  .  |||       |       | |     .    || |        .||
MVSY1    - NPNDVWARPDFGDNYDFPTNFNCKDYPYHGGVSY-GYGNGTP--GYYPDL -521

GL1FULL  - GRPDVAEWWGNNYKKLFSIGLDFVWQDMTVPAMMPHKIGDDINVKPDGNW -626
           | .|   ||| .|   ||  .||.||||||| ||                |            .
MVSY1    - NREEVRIWWGLQYEYLFNMGLEFVWQDMTTPA---------IH----SSY -558

GL1FULL  - PNADDPSNGQYNWKTYHPQVLVTDMRYENHG-REPMVTQRNIHAYTLCES -675
               |    |                .|||       |   :     .       . .| |  .
MVSY1    - ---GDM-KG------LPTRLLVTADSVTNASEKKLAIESWALYSYNLHKA -598

GL1FULL  - TRKE-GIVENADTLTKFRRSYIISRGGYIGNQHFGGMWVGDNSTTSNYIQ -724
           |      | .|         | .| .|. || |  |      | |.| |||..|   .
MVSY1    - TFHGLGRLES----RKNKRNFILGRGSYAGAYRFAGLWTGDNASTWEFWK -644

GL1FULL  - MMIANNINMNMSCLPLVGSDIGGFTSYDNE-NQRTPCTGDLMVRYVQAGC -773
           . ... . . . .  . . ||| |||        |        .  |. .|.|.
MVSY1    - ISVSQVLSLGLNGVCIAGSDTGGFEPARTEIGEEKYCSPELLIRWYTGSF -694

GL1FULL  - LLPWFRNHYDRWIESKDHGKDYQELYMYPNEMDT---L-------R---- -809
           |||| |||| .       ||   |  .|| | ||   ..|    |             .
```

```
MVSY1     - LLPWLRNHYVK----KDR-KWFQEPYAYPKHLETHPELADQAWLYKSVLE -739

GL1FULL   - --KF-VEFRYRWQEVLYTAMYQNAA----FGK-PIIKAA--SMYNNDSNV -849
            .. || ||    .|| |.||        .  ..  .  . . |.|.
MVSY1     - ICRYWVELRYSLIQLLYDCMFQNVVDGMPLARSMLLTDTEDTTFFNESQ- -788

GL1FULL   - RRAQNDHFLLGGHDGYRILCAPVVWENS-TE-----RELYLPVLTQWY-- -891
            .     ..  | |   || ||..   | |    |..|||.   ||
MVSY1     - -KFLDNQYM-AGDD---ILVAPIL--HSRNEVPGENRDVYLPLFHTWYPS -831

GL1FULL   - KFGP-DFDT-----KPLEGAM--NGGDRIY---NYPVPQSESPIFVREGA -930
              | | |       |.||   |    ||     | .     |...||||
MVSY1     - NLRPWD-DQGVALGNPVEGGSVINYTARIVAPEDYNLFHNVVPVYIREGA -880

GL1FULL   - ILPT---RYTLNGENKSLNTYTDEDPLVFEVFPLGNNRADGMCYLDDGGV -977
           |.|   |    ||             |. | ..| |  .      ||||| |
MVSY1     - IIPQIQVR-QWIGE--G-----GPNPIKFNIYP-GKDKEYVT-YLDDG-V -919

GL1FULL   - TTNAEDNGKFSVVKVAAEQD--GGTET---ITF-----TNDCYEYVFGGP -1017
            .  .          . | ||    | .    ..       |||     |
MVSY1     - SRDSAPDD-LPQYREAYEQAKVEGKDVQKQLAVIQGNKTND-----FSAS -963

GL1FULL   - FYVR-VRGAQSPSNI-HVS---SGAGSQDMKVSS-ATSRAA-------LF -1054
              .  .|       |  |   . . . . |      |.         |.
MVSY1     - GIDKEAKGYHRKVSIKQESKDKTRTVTIEPKHNGYDPSKEVGNYYTIILW -1013

GL1FULL   - NDGE-NGDFWVD--QETDSL--WLK---------LPNVVLPDA--VI-T-- -1086
                 |   || | |  .    .            |  ||.     || |
MVSY1     - YAPGFDGSI-VDVSQATVNIEGGVECEIFKNTGLHTVVV-NVKEVIGTTK -1061

GL1FULL   - ---IT      -1088
              ||
MVSY1     - SVKITCTTA -1070
```

^^

4. Sequence Id No. 1 (SIN1) aligned with Sequence Id No. 5 (SIN5)

```
Sequence 1: SIN1, (1088 residues)
Sequence 2: SIN5, (1092 residues)

using the parameters:
Comparison matrix: BLOSUM62
Number of alignments computed: 20
Gap open penalty: 12
Gap extension penalty: 4
```

74.6% identity in 1096 residues overlap; Score: 4455.0; Gap frequency: 1.3%

```
SIN1,              1 MFSTLAFVAPSALGASTFVGAEVRSNVRIHSAFPAVHTATRKTNRLNVSMTALSDKQTAT
SIN5,              1 MFPTLTFIAPSALAASTFVGADIRSGIRIQSALPAVRNAVRRSKHYNVSMTALSDKQTAI
                     ** ** * ***** *******  **  ** ** ***  * *     *************

SIN1,             61 AGSTDNPDGIDYKTYDYVGVWGFSPLSNTNWFAAGSSTPGGITDWTATMNVNFDRIDNPS
SIN5,             61 SIGPDNPDGINYQNYDYIPVAGFTPLSNTNWYAAGSSTPGGITDWTATMNVKFDRIDNPS
                         ****** *  ***  * ** ******* ******************* ********

SIN1,            121 ITVQHPVQVQVTSYNNNSYRVRFNPDGPIRDVTRGPILKQQLDWIRTQELSEGCDPGMTF
SIN5,            121 YSNNHPVQIQVTSYNNNSFRIRFNPDGPIRDVSRGPILKQQLTWIRNQELAQGCNPNMSF
                         **** ********* * *********** ********* *** ***  ** * * *

SIN1,            181 TSEGFLTFETKDLSVIIYGNFKTRVTRKSDGKVIMENDEVGTASSGNKCRGLMFVDRLYG
SIN5,            181 SPEGFLSFETKDLNVIIYGNCKMRVTKK-DGYLVMENDECNSQSDGNKCRGLMYVDRLYG
                       **** ****** ****** * *** * **   *****    * ******** ******

SIN1,            241 NAIASVNKNFRNDAVKQEGFYGAGEVNCKYQD-----TYILERTGIAMTNYNYDNLNYNQ
SIN5,            240 NAIASVQTNFHKDTSRNEKFYGAGEVNCRYEEQGKAPTYVLERSGLAMTNYNYDNLNYNQ
                     ******  **  *    * ********* *       ** *** * **************
```

```
SIN1,         296 WDLRPPHHDGALNPDYYIPMYYAAPWLIVNGCAGTSEQYSYGWFMDNVSQSYMNTGDTTW
SIN5,         300 PDVVPPGYPD--HPNYYIPMYYAAPWLVVQGCAGTSKQYSYGWFMDNVSQSYMNTGDTAW
                   *  **       * ************ * ****** ********************* *

SIN1,         356 NSGQEDLAYMGAQYGPFDQHFVYGAGGGMECVVTAFSLLQGKEFENQVLNKRSVMPPKYV
SIN5,         358 NCGQENLAYMGAQYGPFDQHFVYGDGDGLEDVVKAFSFLQGKEFEDKKLNKRSVMPPKYV
                  * *** ****************** * * * ** *** *******   ************

SIN1,         416 FGFFQGVFGTSSLLRAHMPAGENNISVEEIVEGYQNNNFPFEGLAVDVDMQDNLRVFTTK
SIN5,         418 FGFFQGVFGALSLLKQNLPAGENNISVQEIVEGYQDNDYPFEGLAVDVDMQDDLRVFTTK
                  *********  ***    ********* ******* *  ************* *******

SIN1,         476 GEFWTANRVGTGGDPNNRSVFEWAHDKGLVCQTNITCFLRNDNEGQDYEVNQTLRERQLY
SIN5,         478 PEYWSANMVGEGGDPNNRSVFEWAHDRGLVCQTNVTCFLRNDNSGKPYEVNQTLREKQLY
                   * * ** ** *************** ******* ******** *  ********* ***

SIN1,         536 TKNDSLTGTDFGMTDDGPSDAYIGHLDYGGGVECDALFPDWGRPDVAEWWGNNYKKLFSI
SIN5,         538 TKNDSLNNTDFGTTSDGPGDAYIGHLDYGGGVECDAIFPDWGRPDVAQWWGENYKKLFSI
                  ******  **** * *** ***************** ********** *** ********

SIN1,         596 GLDFVWQDMTVPAMMPHKIGDDINVKPDGN---WPNADDPSNGQYNWKTYHPQVLVTDMR
SIN5,         598 GLDFVWQDMTVPAMMPHRLGDAVNKNSGSSAPGWPNENDPSNGRYNWKSYHPQVLVTDMR
                  *****************  **  *         ***  ***** **** ***********

SIN1,         653 Y-ENHGREPMVTQRNIHAYTLCESTRKEGIVENADTLTKFRRSYIISRGGYIGNQHFGGM
SIN5,         658 YGAEYGREPMVSQRNIHAYTLCESTRREGIVGNADSLTKFRRSYIISRGGYIGNQHFGGM
                  *    ****** ************** **** *** ************************

SIN1,         712 WVGDNSTTSNYIQMMIANNINMNMSCLPLVGSDIGGFTSYDNENQRTPCTGDLMVRYVQA
SIN5,         718 WVGDNSATESYLQMMLANIINMNMSCLPLVGSDIGGFTQYNDAGDPTP--EDLMVRFVQA
                  ****** *  * *** ** ******************* *      **   ***** ***

SIN1,         772 GCLLPWFRNHYDRWIESKDHGKDYQELYMYPNEMDTLRKFVEFRYRWQEVLYTAMYQNAA
SIN5,         776 GCLLPWFRNHYDRWIESKKHGKKYQELYMYPGQKDTLKKFVEFRYRWQEVLYTAMYQNAT
                  ****************** *** ********   *** *********************

SIN1,         832 FGKPIIKAASMYNNDSNVRRAQNDHFLLGGHDGYRILCAPVVWENSTERELYLPVLTQWY
SIN5,         836 TGEPIIKAAPMYNNDVNVYKSQNDHFLLGGHDGYRILCAPVVRENATSREVYLPVYSKWF
                   * ****** ***** **   ********************* ** * ** ****   *

SIN1,         892 KFGPDFDTKPLEGAMNGGDRIYNYPVPQSESPIFVREGAILPTRYTLNGENKSLNTYTDE
SIN5,         896 KFGPDFDTKPLENEIQGGQTLYNYAAPLNDSPIFVREGTILPTRYTLDGVNKSINTYTDN
                  ************    **   ***  *   ******** ******** * *** *****

SIN1,         952 DPLVFEVFPLGNNRADGMCYLDDGGVTTNAEDNGKFSVVKVAAEQDGGTETITFTNDCYE
SIN5,         956 DPLVFELFPLENNQAHGLFYHDDGGVTTNAEDFGKYSVISVKAAQEGSQMSVKFDNEVYE
                  ****** *** ** * *  * *********** ** **  * * * *      * *  **

SIN1,        1012 YVFGGPFYVRVRGAQSPSNIHVSSGAGSQDMKVSSATSRAALFNDGENGDFWVDQETDSL
SIN5,        1016 HQWGASFYVRVRNMGAPSNINVSSQIGQQDMQQSSVSSRAQMFTSANDGEYWVDQSTNSL
                     *  ******    **** ***  * ***  **  ***  *     *  **** * **

SIN1,        1072 WLKLPNVVLPDAVITI
SIN5,        1076 WLKLPGAVIQDAAITV
                  *****  *  ** **

^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
```

5. Sequence Id No. 1 (SIN1) aligned with Sequence Id No. 6 (SIN6)

```
Sequence 1: SIN1, (1088 residues)
Sequence 2: SIN6, (570 residues)

using the parameters:
Comparison matrix: BLOSUM62
Number of alignments computed: 20
Gap open penalty: 12
Gap extension penalty: 4
```

77.9% identity in 578 residues overlap; Score: 2451.0; Gap frequency: 1.9%

```
SIN1,         283 MTNYNYDNLNYNQWDLRPPHHDGALNPDYYIPMYYAAPWLIVNGCAGTSEQYSYGWFMDN
SIN6,           1 MTNYNYDNLNYNQPDLIPPGHDS--DPDYYIPMYFAAPWVIAHGYRGTSDQYSYGWFLDN
                  ************* ** ** **    ******** **** *  *  *** ******* **

SIN1,         343 VSQSYMNTGDTTWNSGQEDLAYMGAQYGPFDQHFVYGAGGGMECVVTAFSLLQGKEFENQ
SIN6,          59 VSQSYTNTGDDAW-AGQKDLAYMGAQCGPFDQHFVYEAGDGLEDVVTAFSYLQGKEYENQ
                  ***** ****  *  ** ******** ********* ** * * ****** ***** ***

SIN1,         403 VLNKRSVMPPKYVFGFFQGVFGTSSLLRAHMPAGENNISVEEIVEGYQNNNFPFEGLAVD
```

```
SIN6,       118 GLNIRSAMPPKYVFGFFQGVFGATSLLRDNLPAGENNVSLEEIVEGYQNQNVPFEGLAVD
                 ** ** **************** ****   ****** * ********* * ********

SIN1,       463 VDMQDNLRVFTTKGEFWTANRVGTGGDPNNRSVFEWAHDKGLVCQTNITCFLRNDNEGQD
SIN6,       178 VDMQDDLRVFTTRPAFWTANKVGEGGDPNNKSVFEWAHDRGLVCQTNVTCFLK--NEKNP
                ***** ******   ***** ** ****** ******** ******* ****   **

SIN1,       523 YEVNQTLRERQLYTKNDSLTGTDFGMTDDGPSDAYIGHLDYGGGVECDALFPDWGRPDVA
SIN6,       236 YEVNQSLREKQLYTKSDSLDNIDFGTTPDGPSDAYIGHLDYGGGVECDALFPDWGRPDVA
                ***** *** ***** ***   *** * ********************************

SIN1,       583 EWWGNNYKKLFSIGLDFVWQDMTVPAMMPHKIGDDINVKPDGN---WPNADDPSNGQYNW
SIN6,       296 QWWGDNYKKLFSIGLDFVWQDMTVPAMMPHRLGDPVGTNSGETAPGWPNDKDPSNGRYNW
                 *** *************************  **            ***  ***** ***

SIN1,       640 KTYHPQVLVTDMRYENHGREPMVTQRNIHAYTLCESTRKEGIVENADTLTKFRRSYIISR
SIN6,       356 KSYHPQVLVTDMRYDDYGRDPIVTQRNLHAYTLCESTRREGIVGNADSLTKFRRSYIISR
                * ************   ** * ***** ********** **** *** ************

SIN1,       700 GGYIGNQHFGGMWVGDNSTTSNYIQMMIANNINMNMSCLPLVGSDIGGFTSYDNENQRTP
SIN6,       416 GGYIGNQHFGGMWVGDNSSTEDYLAMMVINVINMNMSGVPLVGSDIGGFTEHDKRN---P
                ****************** *  *  **  * ******  ***********  *  *   *

SIN1,       760 CTGDLMVRYVQAGCLLPWFRNHYDRWIESKDHGKDYQELYMYPNEMDTLRKFVEFRYRWQ
SIN6,       473 CTPDLMMRFVQAGCLLPWFRNHYDRWIESKKHGKNYQELYMYRDHLDALRSFVELRYRWQ
                ** *** * ********************* *** *******    * ** *** *****

SIN1,       820 EVLYTAMYQNAAFGKPIIKAASMYNNDSNVRRAQNDHF
SIN6,       533 EVLYTAMYQNALNGKPIIKTVSMYNNDMNVKDAQNDHF
                ***********  ******  ****** **  ******
^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
```

```
===03-MAY-2001===================NALIGN========================PC/GENE===

*******************************************
* ALIGNMENT OF TWO NUCLEOTIDE SEQUENCES. *
*******************************************

The two sequences to be aligned are:

SEQIDNO7.
Total number of bases: 3267.

SEQIDNO8.
Total number of bases: 3276.

Open gap cost      : 50
Unit gap cost      : 20

The character to show that two aligned residues are identical is '|'

SEQIDNO7  - ATGTTTTCAACCCTTGCGTTTGTCGCACCTAGTGCGCTGGGAGCCAGTAC -50
            |||| | |||||||  | || || || ||||||||||| || ||||| ||
SEQIDNO8  - ATGTATCCAACCCTCACCTTCGTGGCGCCTAGTGCGCTAGGGGCCAGAAC -50

SEQIDNO7  - CTTCGTAGGGGCGGAGGTC---AGGTCAAATGTTCGTATCCATTCCGCTT -97
             |||    | | ||   |    |||||| |  ||| ||| ||||| | | 
SEQIDNO8  - TTTCACGTGTGTGGGCATTTTTAGGTCACACATTCTTATTCATTCGGTTG -100

SEQIDNO7  - TTCCAGCTGTGCACACAGCTACTCGCAAAACCAATCGCCTCAATGTATCC -147
            ||||||| ||||    ||||   ||||||| ||| |||||||||||||||
SEQIDNO8  - TTCCAGCGGTGCGTCTAGCTGTGCGCAAAAGCAACCGCCTCAATGTATCC -150

SEQIDNO7  - ATGACCGCATTGTCCGACAAACAAACGGCTACTGCGGGTAGTACAGACAA -197
            ||| |||| |||| ||||||||  || |||  | | ||  | |  |||||
SEQIDNO8  - ATGTCCGCTTTGTTCGACAAACCGACTGCTGTTACTGGAGGGAAGGACAA -200

SEQIDNO7  - TCCGGACGGTATCGACTACAAGACCTACGATTACGTCGGAGTATGGGGTT -247
             ||||||  |||| | ||||  || || || ||||||   || ||| | |
SEQIDNO8  - CCCGGACAATATCAATTACACCACTTATGACTACGTCCCTGTGTGGCGCT -250

SEQIDNO7  - TCAGCCCCCTCTCCAACACGAACTGGTTTGCTGCCGGCTCTTCTACCCCG -297
            ||  |||||||  ||| |||||||||||||||||||| ||||| || || 
SEQIDNO8  - TCGACCCCCTCAGCAATACGAACTGGTTTGCTGCCGGATCTTCCACTCCC -300

SEQIDNO7  - GGTGGCATCACTGATTGGACGGCTACAATGAATGTCAACTTCGACCGTAT -347
            || |  ||    || |||||||| ||||||||||| ||||||||||||||
SEQIDNO8  - GGCGATATTGACGACTGGACGGCGACAATGAATGTGAACTTCGACCGTAT -350

SEQIDNO7  - CGACAATCCGTCCATCACTGTCCAGCATCCCGTTCAGGTTCAGGTCACGT -397
            ||||||||| ||| ||||| || || | || |||||||||||||||||||
SEQIDNO8  - CGACAATCCATCCTTCACTCTCGAGAAACCGGTTCAGGTTCAGGTCACGT -400

SEQIDNO7  - CATACAACAACAACAGCTACAGGGTTCGCTTCAACCCTGATGGCCCTATT -447
            ||||||| |||||  | | |||||||||||||||||||||||| ||||||
SEQIDNO8  - CATACAAGAACAATTGTTTCAGGGTTCGCTTCAACCCTGATGGTCCTATT -450
```

```
SEQIDNO7  - CGTGATGTGACTCGTGGGCCTATCCTCAAGCAGCAACTAGATTGGATTCG -497
SEQIDNO8  - CGCGATGTGGATCGTGGGCCTATCCTCCAGCAGCAACTAAATTGGATCCG -500

SEQIDNO7  - AACGCAGGAGCTGTCAGAGGGATGTGATCCCGGAATGACTTTCACATCAG -547
SEQIDNO8  - GAAGCAGGAGCAGTCGAAGGGGTTTGATCCTAAGATGGGCTTCACAAAAG -550

SEQIDNO7  - AAGGTTTCTTGACTTTTGAGACCAAGGATCTAAGCGTCATCATCTACGGA -597
SEQIDNO8  - AAGGTTTCTTGAAATTTGAGACCAAGGATCTGAACGTTATCATATATGGC -600

SEQIDNO7  - AATTTCAAGACCAGAGTTACGAGAAAGTCTGACGGCAAGGTCATCATGGA -647
SEQIDNO8  - AATTTTAAGACTAGAGTTACGAGGAAGAGGGATGGAAAAGGGATCATGGA -650

SEQIDNO7  - AAATGATGAAGTTGGAACTGCATCGTCCGGGAACAAGTGCCGGGGATTGA -697
SEQIDNO8  - GAATAATGAAGTGCCGGCAGGATCGTTAGGGAACAAGTGCCGGGGATTGA -700

SEQIDNO7  - TGTTCGTTGATAGATTATACGGTAACGCTATCGCTTCCGTCAACAAGAAC -747
SEQIDNO8  - TGTTTGTCGACAGGTTGTACGGCACTGCCATCGCTTCCGTTAATGAAAAT -750

SEQIDNO7  - TTCCGCAACGACGCGGTCAAGCAGGAGGGATTCTATGGTGCAGGTGAAGT -797
SEQIDNO8  - TACCGCAACGATCCCGACAGGAAAGAGGGGTTCTATGGTGCAGGAGAAGT -800

SEQIDNO7  - CAACTGTAAGTACCAGGACACC-------------------TACATCTTAG -829
SEQIDNO8  - AAACTGCGAGTTTTGGGACTCCGAACAAAACAGGAACAAGTACATCTTAG -850

SEQIDNO7  - AACGCACTGGAATCGCCATGACAAATTACAACTACGATAACTTGAACTAT -879
SEQIDNO8  - AACGAACTGGAATCGCCATGACAAATTACAATTATGACAACTATAACTAC -900

SEQIDNO7  - AACCAGTGGGACCTTAGACCTCCGCATCATGATGGTGCCCTCAACCCAGA -929
SEQIDNO8  - AACCAGTCAGATCTTATTGCTCCAGGATATCCT------TCCGACCCGAA -944

SEQIDNO7  - CTATTATATTCCAATGTACTACGCAGCACCTTGGTTGATCGTTAATGGAT -979
SEQIDNO8  - CTTCTACATTCCCATGTATTTTGCAGCACCTTGGGTAGTTGTTAAGGGAT -994

SEQIDNO7  - GCGCCGGTACTTCGGAG---CAGTACTCGTATGGATGGTTCATGGACAAT -1026
SEQIDNO8  - GCAGTGGCAACAGCGATGAACAGTACTCGTACGGATGGTTTATGGATAAT -1044

SEQIDNO7  - GTCTCTCAATCTTACATGAATACTGGAGATACTACCTGGAATTCTGGACA -1076
SEQIDNO8  - GTCTCCCAAACTTACATGAATACTGGTGGTACTTCCTGGAACTGTGGAGA -1094

SEQIDNO7  - AGAGGACCTGGCATACATGGGCGCGCAGTATGGACCATTTGACCAACATT -1126
```

```
              ||| || ||||||||||||| || ||||   || |||||||||||||||
SEQIDNO8  - GGAGAACTTGGCATACATGGGAGCACAGTGCGGTCCATTTGACCAACATT -1144

SEQIDNO7  - TTGTTTACGGTGCTGGGGGTGGATGGAATGTGTGGTCACAGCGTTCTCT -1176
            |||| || |||| ||| | |||   | ||    ||| |||  ||||||||||
SEQIDNO8  - TTGTGTATGGTGATGGAGATGGTCTTGAGGATGTTGTCCAAGCGTTCTCT -1194

SEQIDNO7  - CTTCTACAAGGCAAGGAGTTCGAGAACCAAGTTCTCAACAAACGTTCAGT -1226
            ||||| |||||||| ||||| |||||||||||||| ||||||||| | ||
SEQIDNO8  - CTTCTGCAAGGCAAAGAGTTTGAGAACCAAGTTCTGAACAAACGTGCCGT -1244

SEQIDNO7  - AATGCCTCCGAAATACGTCTTTGGTTTCTTCCAGGGTGTTTTCGGGACTT -1276
            |||||||||||||| || ||||||| ||| ||||| || || |||| |
SEQIDNO8  - AATGCCTCCGAAATATGTGTTTGGTTACTTTCAGGGAGTCTTTGGGATTG -1294

SEQIDNO7  - CTTCCTTGTTGAGAGCGCATATGCCAGCAGGTGAGAACAACATCTCAGTC -1326
            |||||||||||||| ||| |   ||||   ||||   || |||||||| ||
SEQIDNO8  - CTTCCTTGTTGAGAGAGCAAAGACCAGAGGGTGGTAATAACATCTCTGTT -1344

SEQIDNO7  - GAAGAAATTGTAGAAGGTTATCAAAACAACAATTTCCCTTTCGAGGGGCT -1376
             |||| ||||| |||||||| |||| ||| || |||||||| |||||| |
SEQIDNO8  - CAAGAGATTGTCGAAGGTTACCAAAGCAATAACTTCCCTTTAGAGGGGTT -1394

SEQIDNO7  - CGCTGTGGACGTGGATATGCAAGACAACTTGCGGGTGTTCACCACGAAGG -1426
             || || || ||||||||||| |  | ||||| ||||||||||||||
SEQIDNO8  - AGCCGTAGATGTGGATATGCAACAAGATTTGCGCGTGTTCACCACGAAGA -1444

SEQIDNO7  - GCGAATTTTGGACCGCAAACAGGGTGGGTACTGGCGGGGATCCAAACAAC -1476
              |||||||||| ||||| | ||| || || || || ||   | || |||
SEQIDNO8  - TTGAATTTTGGACGGCAAATAAGGTAGGCACCGGGGGAGACTCGAATAAC -1494

SEQIDNO7  - CGATCGGTTTTGAATGGGCACATGACAAAGGCCTTGTTTGTCAGACAAA -1526
               ||||| |||||||||||||||||||||||||||| |||||||| ||
SEQIDNO8  - AAGTCGGTGTTTGAATGGGCACATGACAAAGGCCTTGTATGTCAGACGAA -1544

SEQIDNO7  - TATAACTTGCTTCCTGAGGAATGATAACGAGGGGCAAGACTACGAGGTCA -1576
            | | ||||||||| |||| || || ||||   |||  ||| ||||| ||||
SEQIDNO8  - TGTTACTTGCTTCTTGAGAAACGACAACGGCGGGGCAGATTACGAAGTCA -1594

SEQIDNO7  - ATCAGACGTTAAGGGAGAGGCAGTTGTACACGAAGAACGACTCCCTGACG -1626
            ||||||| || ||||||| |    |||||||||||||| ||||| ||||||
SEQIDNO8  - ATCAGACATTGAGGGAGAAGGGTTTGTACACGAAGAATGACTCACTGACG -1644

SEQIDNO7  - GGTACGGATTTTGGAATGACCGACGACGGCCCCAGCGATGCGTACATCGG -1676
               ||   | || ||||  ||| ||||||| || ||||||||||||| ||
SEQIDNO8  - AACACTAACTTCGGAACTACCAACGACGGGCCGAGCGATGCGTACATTGG -1694

SEQIDNO7  - TCATCTGGACTATGGGGGTGGAGTAGAATGTGATGCACTTTTCCCAGACT -1726
             |||||||||||||| || ||||    | |||||||||||||||||||||
SEQIDNO8  - ACATCTGGACTATGGTGGCGGAGGGAATTGTGATGCACTTTTCCCAGACT -1744

SEQIDNO7  - GGGGACGGCCTGACGTGGCCGAATGGTGGGGAAATAACTATAAGAAACTG -1776
            |||| || || |  ||||| |||||||||||  ||||||| |  || ||
SEQIDNO8  - GGGGTCGACCGGGTGTGGCTGAATGGTGGGGTGATAACTACAGCAAGCTC -1794
```

```
SEQIDNO7 - TTCAGCATTGGTCTCGACTTCGTCTGGCAAGACATGACTGTTCCAGCAAT -1826
SEQIDNO8 - TTCAAAATTGGTCTGGATTTCGTCTGGCAAGACATGACAGTTCCAGCTAT -1844

SEQIDNO7 - GATGCCGCACAAAATTGGCGATGACATCAATGTGAAACCGGATGGGAATT -1876
SEQIDNO8 - GATGCCACACAAAGTTGGCGACGCAGTCGATACGAGATCACCTTACGGCT -1894

SEQIDNO7 - GGCCGAATGCGGACGATCCGTCCAATGGACAATACAACTGGAAGACGTAC -1926
SEQIDNO8 - GGCCGAATGAGAATGATCCTTCGAACGGACGATACAATTGGAAATCTTAC -1944

SEQIDNO7 - CATCCCCAAGTGCTTGTAACTGATATGCGTTATGAGAATCATGGTCGGGA -1976
SEQIDNO8 - CATCCACAAGTTCTCGTAACTGATATGCGATATGAGAATCATGGAAGGGA -1994

SEQIDNO7 - ACCGATGGTCACTCAACGCAACATTCATGCGTATACACTGTGCGAGTCTA -2026
SEQIDNO8 - ACCGATGTTCACTCAACGCAATATGCATGCGTACACACTCTGTGAATCTA -2044

SEQIDNO7 - CTAGGAAGGAAGGGATCGTGGAAAACGCAGACACTCTAACGAAGTTCCGC -2076
SEQIDNO8 - CGAGGAAGGAAGGGATTGTTGCAAATGCAGACACTCTAACGAAGTTCCGC -2094

SEQIDNO7 - CGTAGCTACATTATCAGTCGTGGTGGTTACATTGGTAACCAGCATTTCGG -2126
SEQIDNO8 - CGCAGTTATATTATCAGTCGTGGAGGTTACATTGGCAACCAGCATTTTGG -2144

SEQIDNO7 - GGGTATGTGGGTGGGAGACAACTCTACTACATCAAACTACATCCAAATGA -2176
SEQIDNO8 - AGGAATGTGGGTTGGAGACAACTCTTCCTCCCAAAGATACCTCCAAATGA -2194

SEQIDNO7 - TGATTGCCAACAATATTAACATGAATATGTCTTGCTTGCCTCTCGTCGGC -2226
SEQIDNO8 - TGATCGCGAACATCGTCAACATGAACATGTCTTGCCTTCCACTAGTTGGG -2244

SEQIDNO7 - TCCGACATTGGAGGATTCACCTCATACGACAATGAGAATCAGCGAACGCC -2276
SEQIDNO8 - TCCGACATTGGAGGTTTTACTTCGTATGA---------TGGACGAAACGT -2285

SEQIDNO7 - GTGTACCGGGGACTTGATGGTGAGGTATGTGCAGGCGGGCTGCCTGTTGC -2326
SEQIDNO8 - GTGTCCCGGGGATCTAATGGTAAGATTCGTGCAGGCGGGTTGCTTACTAC -2335

SEQIDNO7 - CGTGGTTCAGGAACCACTATGATAGGTGGATCGAGTCCAAGGACCACGGA -2376
SEQIDNO8 - CGTGGTTCAGAAACCACTATGGTAGGTTGGTCGAGGGCAAGCAAGAGGGA -2385

SEQIDNO7 - AAGGACTACCAGGAGCTGTACATGTATCCGAATGAAATGGATACGTTGAG -2426
SEQIDNO8 - AAATACTATCAAGAACTGTACATGTACAAGGACGAGATGGCTACATTGAG -2435

SEQIDNO7 - GAAGTTCGTTGAATTCCGTTATCGCTGGCAGGAAGTGTTGTACACGGCCA -2476
```

```
            || ||| ||||||||||||| |||||||||| |||||||||| || |
SEQIDNO8 - AAAATTCATTGAATTCCGTTACCGCTGGCAGGAGGTGTTGTACACTGCTA -2485

SEQIDNO7 - TGTACCAGAATGCGGCTTTCGGAAAGCCGATTATCAAGGCTGCTTCGATG -2526
           |||||||||||||||||||||| || |||||||||||||| ||||| |||
SEQIDNO8 - TGTACCAGAATGCGGCTTTCGGGAAACCGATTATCAAGGCAGCTTCCATG -2535

SEQIDNO7 - TACAATAACGACTCAAACGTTCGCAGGGCGCAGAACGATCATTTCCTTCT -2576
           ||| | ||||||  |||||||||| | || ||| | || || ||||||||
SEQIDNO8 - TACGACAACGACAGAAACGTTCGCGGCGCACAGGATGACCACTTCCTTCT -2585

SEQIDNO7 - TGGTGGACATGATGGATATCGCATTCTGTGCGCGCCTGTTGTGTGGGAGA -2626
            || ||||| ||||||||||| ||| |||| || ||||||||||||||||
SEQIDNO8 - CGGCGGACACGATGGATATCGTATTTTGTGTGCACCTGTTGTGTGGGAGA -2635

SEQIDNO7 - ATTCGACCGAACGCGAATTGTACTTGCCCGTGCTGACCCAATGGTACAAA -2676
           || | |||   |||||  |||||||||| ||||||||| |||||||||||
SEQIDNO8 - ATACAACCAGTCGCGATCTGTACTTGCCTGTGCTGACCAAATGGTACAAA -2685

SEQIDNO7 - TTCGGTCCCGACTTTGACACCAAGCCTCTGGAAGGAGCGATGAACGGAGG -2726
           ||||| || |||| |||||||||||  |||||    ||| || | |||||
SEQIDNO8 - TTCGGCCCTGACTATGACACCAAGCGCCTGGATTCTGCGTTGGATGGAGG -2735

SEQIDNO7 - GGACCGAATTTACAACTACCCTGTACCGCAAAGTGAATCACCAATCTTCG -2776
           | |    ||| | |||||  ||||  | ||||| || || || || || |
SEQIDNO8 - GCAGATGATTAAGAACTATTCTGTGCCACAAAGCGACTCTCCGATATTTG -2785

SEQIDNO7 - TGAGAGAAGGTGCGATTCTCCCTACCCGCTACACGTTGAACGGTGAAAAC -2826
           |||| ||||| || |||||||||||||||||||||||| |||||   |||
SEQIDNO8 - TGAGGGAAGGAGCTATTCTCCCTACCCGCTACACGTTGGACGGTTCGAAC -2835

SEQIDNO7 - AAATCATTGAACACGTACACGGACGAAGATCCGTTGGTGTTTGAAGTATT -2876
           || ||| ||||||||||||| ||| |||| |||||||||||||| |||||
SEQIDNO8 - AAGTCAATGAACACGTACACAGACAAAGACCCGTTGGTGTTTGAGGTATT -2885

SEQIDNO7 - CCCCCTCGGAAACAACCGTGCCGACGGTATGTGTTATCTTGATGATGGCG -2926
           ||| || |||||||||||||||||||||||||||||||||||||||||||
SEQIDNO8 - CCCTCTTGGAAACAACCGTGCCGACGGTATGTGTTATCTTGATGATGGCG -2935

SEQIDNO7 - GTGTGACCACCAATGCTGAAGACAATGGCAAGTTCTCTGTCGTCAAGGTG -2976
           || | || ||  ||||||| ||| ||||||| ||||||||  |||| || 
SEQIDNO8 - GTATTACTACAGATGCTGAGGACCATGGCAAATTCTCTGTTATCAATGTC -2985

SEQIDNO7 - GCAGCGGAGCAGGATGGTGGTACGGAGACGATAACGTTTACGAATGATTG -3026
           | |||    | | | |||| ||||  |||||| | |||| || ||||   
SEQIDNO8 - GAAGCCTTACGGAAAGGTGTTACGACGACGATCAAGTTTGCGTATGACAC -3035

SEQIDNO7 - CTATGAGTACGTTTTCGGTGGACCGTTCTACGTTCGAGTGCGCGGCGCTC -3076
            ||| | ||||| || | ||| || |||||||||||| | ||      | 
SEQIDNO8 - TTATCAATACGTATTTGATGGTCCATTCTACGTTCGAATCCGTAATCTTA -3085

SEQIDNO7 - AGTCGCCGTCGAACATCCACGTGTCTTCTGGAGCGGGTTCTCAGGACATG -3126
            | |  | || || ||  ||||||||||||||||||||    ||||||||
SEQIDNO8 - CGACTGCATCAAAAATTAACGTGTCTTCTGGAGCGGGTGAAGAGGACATG -3135
```

```
SEQIDNO7   - AAGGTGAGCTCTGCCACTTCCAGGGCTGCGCTGTTCAATGACGGGGAGAA -3176
             |    || |||||| |  || ||||| ||  |||||| ||| || |
SEQIDNO8   - ACACCGACCTCTGCGAACTCGAGGGCAGCTTTGTTCAGTGATGGAGGTGT -3185

SEQIDNO7   - CGGTGATTTCTGGGTTGACCAGGAGACAGATTCTCTGTGGCTGAAGTTGC -3226
              || || | ||||| |||| | || ||   |||||||||| |||||||||
SEQIDNO8   - TGGAGAATACTGGGCTGACAATGATACGTCTTCTCTGTGGATGAAGTTGC -3235

SEQIDNO7   - CCAACGTTGTTCTCCCGGACGCTGTGATCACAATTACCTAA -3267
             | ||| | ||||| |  ||||||||||| || ||||| ||
SEQIDNO8   - CAAACCTGGTTCTGCAAGACGCTGTGATTACCATTACGTAG -3276
```

```
Identity  : 2550 (78.05%)
Number of gaps inserted in SEQIDNO7: 3
Number of gaps inserted in SEQIDNO8: 2

===03-MAY-2001===================NALIGN========================PC/GENE===
```

===03-MAY-2001===================NALIGN=========================PC/GENE===

```
******************************************
* ALIGNMENT OF TWO NUCLEOTIDE SEQUENCES. *
******************************************
```

The two sequences to be aligned are:

SEQIDNO7.
Total number of bases: 3267.

SEQIDNO9.
Total number of bases: 3201.

```
Open gap cost      : 50
Unit gap cost      : 20
```

The character to show that two aligned residues are identical is '|'

```
SEQIDNO7  - ATGTTTTCAACCCTTGCGTTTGTCGCACCTAGTGCGCTGGGAGCCAGTAC -50
            |||     |     ||    | ||     | |       | || |
SEQIDNO9  - ATGGCAGGATTTTCTGATCCTCTCAACTTTGCAAAGCAGAAGACTACTA -50

SEQIDNO7  - CTTCGTAGGGGCGGAGGTCAGGTCAAATGTTCGTATCCATTCCGCTTTTC -100
            |   || | |   ||    | |     |      |||  |   |
SEQIDNO9  - CAGTGTTGCGCTAGACTGGAAGGGCCCTCAAAAAATCATTGGAGTAGACA -100

SEQIDNO7  - CAGCTGTGCACACAGCTACTCGCAAAACCAATCGCCTCAATGTATCCATG -150
            |  ||   |  |     ||        ||||   |    ||| |    ||
SEQIDNO9  - CTACTCCTCCAAAGAGCACCAAGTTCCCCAAAAACTGGCATGGAG---TG -147

SEQIDNO7  - ACCGCATTGTCCGACAAACAAACGGCTACTGCGGGTAGTACAGACAATCC -200
            | |      | |||  |    ||      || || |         |  ||
SEQIDNO9  - AACTTGAGATTCGATGATGGGACTTTAGGTGTGGTTCAGTTCATTAGGCC -197

SEQIDNO7  - GGACGGTATCGACTACAAGACCTACGATTACGTCGGAGTATGGGGTTTCA -250
            | | | |         |    | ||  |    ||      |  |
SEQIDNO9  - GTGCGTTTGGAGGGTTAGATACGACCCTGGTTTCAAGACCTCTGACGAGT -247

SEQIDNO7  - GCCCCCTCTCCAACACGAACTGGTTTGCTGCCGGCTCTTCTACCCCGGGT -300
                       || ||||      |||     |  | |   |   |
SEQIDNO9  - ATGGTGATGAGAATACGAGGACAATTGTGCAAGATTATATGAGTACTCTG -297

SEQIDNO7  - GGCATCACTGATTGGACGGCTACAATGAATGTCAACTTCGACCGTATCGA -350
             | |  |        ||   || |     |           |      ||
SEQIDNO9  - AGTAATAAATTGGATACTTATAGAGGTCTTACGTGGGAAACCAAGTGTGA -347

SEQIDNO7  - CAATCCGTCCATCACTGTCCAGCATCCCGTTCAGGTTCAGGTCACGTCAT -400
              || ||   |    | |  | | || | | || | |||   |   | |
SEQIDNO9  - GGATTCGGG-AGATTTCTTTACCTTCTCATCCAAGGTCACCGCCGTTGAA -396

SEQIDNO7  - ACAACAACAACAGCTACAGGGTTCGCTTCAACCCTGATGGCCCTATTCGT -450
            | | |  | |  |   | |        ||  |  ||    |   ||||
SEQIDNO9  - AAATC--CGAGCGGACCCGCAACAAGGTCGGCGATGGCCTCAGAATTCAC -444
```

```
SEQIDNO7 - GATGTGACTCGTGGGCCTATCCTCAAGCAGCAACTAGATTGGATTCGAAC -500

SEQIDNO9 - CTATGGAAAAGCCCTTTCCGCATCCAAGTAGTGCGCACCTTGACCCCTTT -494
```

```
SEQIDNO7 - GCAGGAGCTGTCAGAGGGATGTGATCCCGGAATGACTTTCACATCAGAAG -550

SEQIDNO9 - GAAGGATCCTTACCCCATTCCAAATGTAGCCGCAGCCGAAGCCCGTGTGT -544

SEQIDNO7 - GTTTCTTGACTTTTGAGACCAAGGATCTAAGCGTCATCATCTACGGAAAT -600

SEQIDNO9 - CCGACAAGGTCGTTTGGCAAACGTCTCCCAAGACATTCAGAAAGAACCTG -594

SEQIDNO7 - TTCAAGACCAGAGTTACGAGAAAGTCTGACGGCAAGGTCATCATGGAAAA -650

SEQIDNO9 - CATCCGCAACACAAGATGCTAAAGGATACAGTTCTTGACATTGTCAAACC -644

SEQIDNO7 - TGATGAAGTTGGAACTGCATCGTCCGGGAACAAGTGCCGGGGATTGATGT -700

SEQIDNO9 - TGGACATGGCGAGTATGTGGGGTGGGGAGAGATGGGAGGTATCCAGTTTA -694

SEQIDNO7 - TCGTTGATAGATTATACGGTAACGCTATCGCTTCCGTCAACAAGAACTTC -750

SEQIDNO9 - TGAAGGAGCCAACATTCATGAACTATTTTAACTTCGACAATATGCAATAC -744

SEQIDNO7 - CGCAACGACGCGGTCAAGCAGGAGGGATTCTATGGTGCAGGTGAAGTCAA -800

SEQIDNO9 - CAGCAAGTCTATGCCCAAGGTGCTCTCGATTCTCGCGAG--------CCA -786

SEQIDNO7 - CTGTAAGTACCAGGACACCTACATCTTAGAACGCACTGGAATCGCCATGA -850

SEQIDNO9 - CTGTACCACTCGGATCCCTTCTATCTTGATGTGAACTCCAACCCGGAGCA -836

SEQIDNO7 - CAAATTACAACTACGATAACTTGAACTATAACCAGTGGGACCTTAGACCT -900

SEQIDNO9 - CAAGA-ATATCACGGCAACCTTTATCGATAACTACTCTCAAATT-GCCAT -884

SEQIDNO7 - CCGCATCATGATGGTGCCCTCAACCCAGACTATTATATTCCAATGTACTA -950

SEQIDNO9 - CGACTTTGGAAAGACCAACTCAGGCTACATCAAGCTGGGAACCAGGTATG -934

SEQIDNO7 - CGCAGCACCTTGGTTGATCGTTAATGGATGCGCCGGTACTTCGGAGCAGT -1000

SEQIDNO9 - GTGGTATCGATTGTTACGGTATCAGTGCGGATACGGTCCCG-GAAATTGT -983

SEQIDNO7 - ACTCGTATGGATGGTTCATGGACAATGTCTCTCAATCTTACATGAATACT -1050

SEQIDNO9 - ACGACTTTATACAGGTCTTGTTGGACGTTCAAAGTTGAAGCCCAGATATA -1033

SEQIDNO7 - GGAGATACTACCTGGAATTCTGGACAAGAGGACCTGGCATACATGGGCGC -1100

SEQIDNO9 - TTCTCGGGGCCCATCAAGCCTGTT-ATGGATACCAACAGGAAAGTGACTT -1082

SEQIDNO7 - GCAGTATGGACCATTTGACCAACATTTTGTTTACGGTGCTGGGGGTGGGA -1150
```

```
SEQIDNO9  - GTATTCTGTGGTCCAGCAGTACCGTGACTGTAAATTTCCACTTGACGGGA -1132

SEQIDNO7  - TGGAATGTGTGGTCACAGCGTTCTCTCTTCTACAAGGCAAGGAGTTCGAG -1200

SEQIDNO9  - TTCACGTCGATGTCGATGTTCAGGACGGCTTCAGAACTTTCACCACCAAC -1182

SEQIDNO7  - AACCAAGTTCTCAACAAACGTTCAGTAATGCCTCCGAAATACGTCTTTGG -1250

SEQIDNO9  - CCACACACTTTCCCTAACCCCAAAGAGATGTTTACTAACTTGAGGAATAA -1232

SEQIDNO7  - TTTCTTCCAGGGTGTTTTCGGGACTTCTTCCTTGTTGAGAGCGCATATGC -1300

SEQIDNO9  - TGGAATCAAGTGCTCCACCAATATCACTCCTGTTATCAGCATTAACAACA -1282

SEQIDNO7  - CAGCAGGTGAGAACAACATCTCAGTCGAAGAAATTGTAGAAGGTTATCAA -1350

SEQIDNO9  - GAGAGGGTGGATACAGTACCCTCCTTGAGGGAGTTGACAAAAAATACTTT -1332

SEQIDNO7  - AACAACAATTTCCCTTTCGAGGGGCTCGCTGTGGACGTGGATATGCAAGA -1400

SEQIDNO9  - ATCATGGACGACAGATATACCGAGGGAACAAGTGGGAATGCGAAGGATGT -1382

SEQIDNO7  - CAACTTGCGGGTGTTCACCACGAAGGGCGAATTTTGGACCGCAAACAGGG -1450

SEQIDNO9  - TCGGTACATGTACTACGGTGGTGGTAATAAGGTTGAGGTCGATCCTAATG -1432

SEQIDNO7  - TGGGTACTGGCGGGGATCCAAACAACCGATCGGTTTTTGAATGGGCACAT -1500

SEQIDNO9  - ATGTTAATGGTCGGCCAGACTTTAAAGACAACTATGACTTCCCCGCGAAC -1482

SEQIDNO7  - GACAAAGGCCTTGTTTGTCAGACAAATATAACTTGCTTCCTGAGGAATGA -1550

SEQIDNO9  - TTCAACAGCAAACAATACCCCTATCATGGTGGTGTGAGCTACGGTTATGG -1532

SEQIDNO7  - TAACGAGGGGCAAGACTACGAGGTCAATCAGACGTTAAGGGAGAGGCAGT -1600

SEQIDNO9  - GAACGGTAGTGCAGGTTTTTACCCGGACCTCAACAGAAAGGAGGTTCGTA -1582

SEQIDNO7  - TGTACACGAAGAACGACTCCCTGACGGGTACGGATTTTGGAATGACCGAC -1650

SEQIDNO9  - TCTGGTGGGGAATGCAGTACAAGTATCTCTTCGATATGGGACTGGAATTT -1632

SEQIDNO7  - GACGGCCCCAGCGATGCGTACATCGGTCATCTGGACTATGGGGGTGGAGT -1700

SEQIDNO9  - GTGTGGCA-AGACATGACTACCCCAGCAATCCACACATCATATGGAGACA -1681

SEQIDNO7  - AGAATGTGATGCACTTTTCCCAGACTGGGGACGGCCTGACGTGGCCGAAT -1750

SEQIDNO9  - TGAAAGGGTTGCCCACCCGTCTACTCGTCACCTCAGACTCCGTCACCAAT -1731

SEQIDNO7  - GGTGGGGAAATAACTATAAGAAACTGTTCAGCATTGGTCTCGACTTCGTC -1800

SEQIDNO9  - GCCTCTGAGAAAAAGCTCGCAATTGAAACTTGGGCTCTCTACTCCTACAA -1781
```

```
SEQIDNO7  - TGGCAAGACATGACTGTTCCAGCAATGATGCCGCACAAAATTGGCGATGA -1850

SEQIDNO9  - TCTCCACAAAGCAACTTGGCATGGTCTTAGTCGTCTCGAATCTCGTAAGA -1831
```

```
SEQIDNO7  - CATCAATGTGAAACCGGATGGGAATTGGCCGAATGCGGACGATCCGTCCA -1900

SEQIDNO9  - ACAAACGAAACTTCATCCTCGGGCGTGGAAG--TTATGCCGGAGCCTATC -1879

SEQIDNO7  - ATGGACAATACAACTGGAAGACGTACCATCCCCAAGTGCTTGTAACTGAT -1950

SEQIDNO9  - GTTTTGCTGGTCTCTGGACTGGGGATAATGCAAGTAACTGGGAATTCTGG -1929

SEQIDNO7  - ATGCGTTATGAGAATCATGGTCGGGAACCGATGGTCACTCAACGCAACAT -2000

SEQIDNO9  - AAGATATCGGTCTCTCAAGTTCTTTCTCTGGGCCTCAATGGTGTGTGCAT -1979

SEQIDNO7  - TCATGCGTATACACTGTGCGAGTCTACTAGGAAGGAAGGGATCGTGGAAA -2050

SEQIDNO9  - CGCGGGGTCTGATACGGGTGGTTTT-----GAACCCTACCGTGATGCAAA -2024

SEQIDNO7  - ACGCAGACACTCTAACGAAGTTCCGCCGTAGCTACATTATCAGTCGTGGT -2100

SEQIDNO9  - TGGGGTCGAGGAGAAATACTGTAGCCCAGAGCTAC-TCATCAGGTGGTAT -2073

SEQIDNO7  - GGTTACATTGGTAACCAGCATTTCGGGGGTATGTGGGTGGGAGACAACTC -2150

SEQIDNO9  - ACTGGTTCATTCCTCTTGCCGTGGCTCAGGAACCATTATGTCAAAAAGGA -2123

SEQIDNO7  - TACTACATCAAACTACATCCAAATGATGATTGCCAACAATATTAACATGA -2200

SEQIDNO9  - CAGGAAATGGTTCCAGGAACCATACTCGTACCCCAAGCATCTTGAAACCC -2173

SEQIDNO7  - ATATGTCTTGCTTGCCTCTCGTCGGCTCCGACATTGGAGGATTCACCTCA -2250

SEQIDNO9  - ATCCAGAACTCGCAGACCAAGCATGGCTCTATAAATCCGTTTTGGAGATC -2223

SEQIDNO7  - TACGACAATGAGAATCAGCGAACGCCGTGTACCGGGGACTTGATGGTGAG -2300

SEQIDNO9  - TGTAGGTACTATGTGGAGCTTAGATACTCCCTCATCCAACTACTTTACGA -2273

SEQIDNO7  - GTATGTGCAGGCGGGCTGCCTGTTGCCGTGGTTCAGGAACCACTATGATA -2350

SEQIDNO9  - CTGCATGTTTCAAAACGTAGTCGACGGTATGCCAATCACCAGATCTATGC -2323

SEQIDNO7  - GGTGGATCGAGTCCAAGGACCACGGAAAGGACTACCAGGAGCTGTACATG -2400

SEQIDNO9  - TCTTGACCGATACTGAGGATACCACCTTCTTCAACGAGAGCCAAAAGTTC -2373

SEQIDNO7  - TATCCGAATGAA-ATGGATACGTTGAGGAAGTTCGTTGAATTCCGTTATC -2449

SEQIDNO9  - CTCGACAACCAATATATGGCTGGTGACGACATTCTTGTTGCACCCATCCT -2423

SEQIDNO7  - GCTGGCAGGAAGTGTTGTACACGGCCATGTACCAGAATGCGGCTTTCGGA -2499
```

```
SEQIDNO9  - CCACAGTCGCAAAGAAATTCCAGGCGAAAACAGAGATGTCTATCTCCCTC -2473

SEQIDNO7  - AAGCCGATTATCAAGGCTGCTTCGATGTACAATAACGACTCAAACGTTCG -2549
SEQIDNO9  - TTTACCACACCTGGTACCCCTCAAATTTGAGACCATGGGACGATCAAG-G -2522

SEQIDNO7  - CAGGGCGCAGAACGATCATTTCCTTCTTGGTGGACATGATGGATATCGCA -2599
SEQIDNO9  - AGTCGCTTTGGGGAATCCTGTCGAAGGTGGTAGTGTCATCAATTATACTG -2572

SEQIDNO7  - TTCTGTGCGCGCCTGTTGTGTGGGAGAATTCGACCGAACGCGAATTGTAC -2649
SEQIDNO9  - CTAGGATTGTTGCACCCGAGGATTATAATCTCTTCCACAGCG--TGGTAC -2620

SEQIDNO7  - TTGCCCGTGCTGACCCAATGGTACAAATTCGGTCCCGACTTTGACACCAA -2699
SEQIDNO9  - CAGTCTACGTTAGAGAGGGTGCCATCATCCCGCAAATCGAAGTACGCCAA -2670

SEQIDNO7  - GCCTCTGGAAGGAGCGATGAACGGAGGGGACCGAATTTACAACTACCCTG -2749
SEQIDNO9  - TGGACTGGCCAGGGGGGAGC-CAACCGCATCAAGTTCAACATCTACCCTG -2719

SEQIDNO7  - TACCGCAAAGTGAATCACCAATCTTCGTGAGAGAAGGTGCGATTCTCCCT -2799
SEQIDNO9  - GAAAGGATAAGGAGTACTGTACCTATCTTGATGATGGTGTTAGCCGTGAT -2769

SEQIDNO7  - ACCCGCTACACGTTGAACGGTGAAAACAAATCATTGAACACGTACACGGA -2849
SEQIDNO9  - AGTGCGCC------------GGAAGACCTCCCACAGTACAAAGAGACCCA -2807

SEQIDNO7  - CGAAGATCCGTTGGTGTTTGAAGTATTCCCCCTCGGAAACAACCGTGCCG -2899
SEQIDNO9  - CGAACA---GTCGAAGGTTGAAGGCGCGGAAATCGCAAAGCAGATTGGAA -2854

SEQIDNO7  - ACGGTATGTGTTATCTTGATGATGGCGGTGTGACCACCAATGCTGAAGAC -2949
SEQIDNO9  - AGAAGACGGGTTACAACATCTCAGGA---------ACCGACCCAGAAGCA -2895

SEQIDNO7  - AATGGCAAGTTCTCTGTCGTCAAGGTGGCAGCGGAGCAGGATGGTGGTAC -2999
SEQIDNO9  - AAGGGTTATCACCGCAAAGTTGCTGTCACACAAACGTCAAAAGACAAGAC -2945

SEQIDNO7  - GGAGACGATAACGTTTACGAATGATTGCTATGAGTACGTTTTCGGTGGAC -3049
SEQIDNO9  - GCGTACTGTCACTATTGAGCCAAAACACAATGGATACGACCCTTCCAAAG -2995

SEQIDNO7  - CGTTCTACGTTCGAGTGCGCGGCGCTCAGTCGCCGTCGAACATCCACGTG -3099
SEQIDNO9  - AGGTGGGTGATTATTATACCATCATTCTTTGGTACGCA-----CCAGGTT -3040

SEQIDNO7  - TCTTCTGGAGCGGGTTCTCAGGACATGAAGGTGAGCTCTGCCACTTCCAG -3149
SEQIDNO9  - TCGATGGCAGCATCGTC----GATGTGAGCAAGACGACTGTGAATGTTGA -3086
```

```
SEQIDNO7  - GGCTGCGCTGTTCAATGACGGGGAGAACGGTGATTTCTGGGTTGACCAGG -3199
            || || | ||    |  | |   | ||          |    |      |
SEQIDNO9  - GGGTGGGGTGGAGCACCAAGTTTATAAGAACTCCGATTTACATACGGTTG -3136
```

```
SEQIDNO7  - AGACAGATTCTCTGTGGCTGAAGTTGCCCAACGTTGTTCTCCCGGACGCT -3249
              |  ||      |  | |||      |||         ||  |  | |
SEQIDNO9  - TTATCGACGTGAAGGAGGTGATCGGTACCACAAAGAGCGTCAAGATCACA -3186

SEQIDNO7  - GTGATCACAATTACCTAA -3267
                  ||     | |||
SEQIDNO9  - TG---TACTGCCGCTTAA -3201
```

```
Identity  : 1108 (34.61%)
Number of gaps inserted in SEQIDNO7: 1
Number of gaps inserted in SEQIDNO9: 21

===03-MAY-2001===================NALIGN========================PC/GENE===
```

===03-MAY-2001===================NALIGN=========================PC/GENE===

```
*******************************************
* ALIGNMENT OF TWO NUCLEOTIDE SEQUENCES. *
*******************************************
```

The two sequences to be aligned are:

SEQIDNO7.
Total number of bases: 3267.

SEQIDNO10.
Total number of bases: 3213.

```
Open gap cost      : 50
Unit gap cost      : 20
```

The character to show that two aligned residues are identical is '|'

```
SEQIDNO7  - ATGTTTTCAACCCTTGCGTTTGTCGCACCTAGTGCGCTGGGAGCCAGTAC -50
            |||     |     | ||    || ||  |  |||   |      | |||
SEQIDNO10 - ATGGCAGGATTA--TCCGACCCTCTCAATTTCTGCAAAGCAGAGGACTAC -48

SEQIDNO7  - CTTCGTAGGGGCGGAGGTCAGGTCAAATGTTCGTATCCATTCCGCTTTTC -100
                 |   | |  |||   |        |  | |     | |||| |
SEQIDNO10 - TACGCTGCTGCCAAAGGCTGGAGTGGCCCTCAGAAGATCATTCGCTATGA -98

SEQIDNO7  - CAGCTGTGCACACAGCTACTCGCAAAACCAATCGCCTCAATGTATCCATG -150
            |        || |  |  |  |  | |  |    |   ||  |  |
SEQIDNO10 - CCA-----GACCCCTCCTCAGGGTACAAAAGATCCGAAAAGCTGGCATGC -143

SEQIDNO7  - ACCGCATTGTCCGACAAACAAACGGCTACTGCGGGTAGTACAGACAATCC -200
                 |   |||     |  |  ||    || | ||||| ||     ||
SEQIDNO10 - GGTAAACCTTCCTTTCGATGACGGGACTATGTGTGTAGTGCAATTCGTCA -193

SEQIDNO7  - GGACGGTATCGACTACAAGACCTACGATTACGTCGGAGTATGGGGTTTCA -250
            |  |      |  |  | |       || ||  | | ||   |  ||
SEQIDNO10 - GACCCTGT--GTTTGGAGGGTTAGATATGACCCCAGTGTCAAGACTTCTG -241

SEQIDNO7  - GCCCCCTCTCCAACACGAACTGGTTTGCTGCCGGCTCTTCTACCCCGGGT -300
                   |  | |   |||    |  |  | |   | | | |  |    |
SEQIDNO10 - ATGAGTACGGCGATGAGAA----TACGAGGACTATTGTACAAGACTACAT -287

SEQIDNO7  - GGCATCACTGATTGGACGGCTACAATGAATGTCAACTTCGACCGTATCGA -350
            | |  | ||| |||||    |  |     |   |  |   ||
SEQIDNO10 - GACTACTCTGGTTGGAAACTTGGACATTTTCAGAGGTCTTACGTGGGTTT -337

SEQIDNO7  - CAATCCGTCCATCACTGTCCAGCATCCCGTTCAGGTTCAGGTCACGTCAT -400
            | |              ||  ||      | ||  |||| |||
SEQIDNO10 - CTACGTTGGAGGA---TTCGGGCGAGTACTACACCTTCAAGTCCGAAGTC -384

SEQIDNO7  - ACAACAACAACAGCTACAGGGTTCGCTTCAACCCTGATGGCCCTATTCGT -450
            ||  |       |  || |      ||  || |  | | | |     | |
SEQIDNO10 - ACTGCCGTGGACGAAACCGAACGGACTCGAAACAAGGTCGGCGACGGCCT -434
```

```
SEQIDNO7  - GATGTGACTCGTGGGCCTATCCTCAAGCAGCAACTAGATTGGATTCGAAC -500

SEQIDNO10 - CAAGATTTACCTATGGAAAAATCCCTTTCGCATCCAG---GTAGTGCGTC -481
```

```
SEQIDNO7  - GCAGGAGCTGTCAGAGGGATGTGATCCCGGAATGACTTTCACATCAGAAG -550

SEQIDNO10 - TCTTGACCCCCCTGGTGGACCCTTTCCCC--ATTCCCAACGTAGCCAATG -529

SEQIDNO7  - GTTTCTTGACTTTTGAGACCAAGGATCTAAGCGTCATCATCTACGGAAAT -600

SEQIDNO10 - CCACAGCCCGTGTGGCCGACAAGGTTGTTTGGCAGACGTCCCCGAAGACG -579

SEQIDNO7  - TTCAAGACCAGAGTTACGAGAAAGTCTGACGGCAAGGTCATCATGGAAAA -650

SEQIDNO10 - TTCAGGAAAAACTTGCATCCGCAGCATAAGATGTTGAAGGATACAGTTCT -629

SEQIDNO7  - TGATGAAGTTGGAACTGCATCGTCCGGGAACAAGTGCCGGGGATTGATGT -700

SEQIDNO10 - TGATATTATCAAGCCGGGGCACGGAGAGTATGTGGGTTGGGGAGAGATGG -679

SEQIDNO7  - TCGTTGATAGATTATACGGTAACGCTATCGCTTCCGTCAACAAGAACTTC -750

SEQIDNO10 - GAGGCATCGAGTTTATGAAGGAGCCAACATTCATGAATTATTTCAACTTT -729

SEQIDNO7  - CGCAACGACGCGGTCAAGCAGGAGGGATTCTATGGTGCAGGTGAAGTCAA -800

SEQIDNO10 - GACAATATGCAATATCAGCAGGTCTATGCACAAGGCGCTCTTGATAGTCG -779

SEQIDNO7  - CTGTAAGTACCAGGACACCTACATCTTAGAACGCACTGGAATCGCCATGA -850

SEQIDNO10 - TGAGCCGTTGTATCACTCTGATCCCTTCTATCTCGACGTGAACTCCAACC -829

SEQIDNO7  - CAAATTACAACTACGATAACTTGAACTATAACCAGTGGGACCTTAGACCT -900

SEQIDNO10 - CAGAGCACAAGAACATTACGGCAACCTTTATCGATAACTACTCTCAGATT -879

SEQIDNO7  - CCGCATCATGATGGTGCCCTCAACCCAGACTATTATATTCCAATGTACTA -950

SEQIDNO10 - GCCATCGACTTTGGGAAGACCAACTCAGGCTAC-ATCAAGCTGGGTACCA -928

SEQIDNO7  - CGCAGCACCTTGGTTGATCGTTAATGGATGCGCCGGTACTTCGGAGCAGT -1000

SEQIDNO10 - GGTATGGCGGTA-TCGATTGTTACGGTATCAGCGCGGATACGGTCCCGGA -977

SEQIDNO7  - ACTCGTATGGATGGTTCATGGACAATGTCTCTCAATCTTACATGAATACT -1050

SEQIDNO10 - GATTGTGCGACTTTATACTGGACT-TGTTGGGCGTTCGAAGTTGAAGCCC -1026

SEQIDNO7  - GGAGATACTACCTGGAATTCTGGACAAGAGGACCTGGCATACATGGCGCC -1100

SEQIDNO10 - AGGTATATTCTCGGAGCCCACCAAGCTTGTTATGGATACCAGCAGGAAAG -1076

SEQIDNO7  - GCAGTATGGACCATTTGACCAACATTTTGTTTACGGTGCTGGGGGTGGGA -1150
```

```
SEQIDNO10 - TGACTTGCATGCTGTTGTTCAGCAGTACCGTGACAC----CAAGTTTCCG -1122

SEQIDNO7  - TGGAATGTGTGGTCACAGCGTTCTCTCTTCTACAAGGCAAGGAGTTCGAG -1200

SEQIDNO10 - CTTGATGGGTTGCATGTCGATGTCGACTTTCAGGACAATTTCAGAACGTT -1172

SEQIDNO7  - AACCAAGTTCTCAACAAACGTTCAGTAATGCCTCCGAAATACGTCTTTGG -1250

SEQIDNO10 - TACCACTAACCCGATTA-CGTTCCCTAATCCCAAAGAAATGTTTACCAA- -1220

SEQIDNO7  - TTTCTTCCAGGGTGTTTTCGGGACTTCTTCCTTGTTGAGAGCGCATATGC -1300

SEQIDNO10 - TCTAAGGAACAATGGAATCAAGTGTTCCACCAACATCACCCCTGTTATCA -1270

SEQIDNO7  - CAGCAGGTGAGAACAACATCTCAGTCGAAGAAATTGTAGAAGGTTATCAA -1350

SEQIDNO10 - GTATCAGAGATCGCCCGAATGGGTACAGTACCCTCAATGAGGGATATGAT -1320

SEQIDNO7  - AACAACAATTTCCCTTTCGAGGGGCTCGCTGTGGACGTGGATATGCAAGA -1400

SEQIDNO10 - AAAAAGTACTTCATCATGGATGACAGATATACCGAGGGGACAAGTGGGGA -1370

SEQIDNO7  - CAACTTGCGGGTGTTCACCACGAAGGGCGAATTTTGGACCGCAAACAGGG -1450

SEQIDNO10 - CCCGCAAAATGTTCGATACTCTTTTTACGGCGGTGGGAACCCGGTTGAGG -1420

SEQIDNO7  - TGGGTACTGGCGGGGATCCAAACAACCGATCGGTTTTTGAATGGGCACAT -1500

SEQIDNO10 - TTAACCCTAATGATGTTTGGGCTCGGCCA----------GACTTTGGAGAC -1461

SEQIDNO7  - GACAAAGGCCTTGTTTGTCAGACAAATATAACTTGCTTCCTGAGGAATGA -1550

SEQIDNO10 - AATTATGACTTCCCTACGAACTTCAACTGCAAAGACTACCCCTATCATGG -1511

SEQIDNO7  - TAACGAGGGGCAAGACTACGAGGTCAATCAGACGTTAAGGGAGAGGCAGT -1600

SEQIDNO10 - TGGTGTGAGTTACGGATATGGGAATGGCACTCCAGGTTACTACCCTGACC -1561

SEQIDNO7  - TGTACACGAAGAACGACTCCCTGACGGGTACGGATTTTGGAATGACCGAC -1650

SEQIDNO10 - TTAACAGAGAGGAGG--TTCGTATCTGGTGGGGATTGCAGTACGAGTATC -1609

SEQIDNO7  - GACGGCCCCAGCGATGCGTACATCGGTCATCTGGACTATGGGGGTGGAGT -1700

SEQIDNO10 - TCTTCAATATGGGACTAGAGTTTGTATGGCAAGATATGACAACCCCAGCG -1659

SEQIDNO7  - AGAATGTGATGCACTTTTCCCAGACTGGGGACGGCCTGACGTGGCCGAAT -1750

SEQIDNO10 - ATCCATTCATCATATGGAGACATGAAAGGGTTGCCCACCCGTCTGCTCGT -1709

SEQIDNO7  - GGTGGGGAAATAACTATAAGAAACTGTTCAGCATTGGTCTCGACTTCGTC -1800

SEQIDNO10 - CACCGCCGACTCAGTTACCAATGCCTCTGAGAAAAAGCTCGCAATTGAAA -1759
```

```
SEQIDNO7  - TGGCAAGACATGACTGTTCCAGCAATGATGCCGCACAAAATTGGCGATGA -1850

SEQIDNO10 - GTTGGGCTCTTTACTCCTACAACCTCCATAAAGCAACCTTCCACGGTCTT -1809
```

```
SEQIDNO7  - CATCAATGTGAAACCGGATGGGAATTGGCCGAATGCGGACGATCCGTCCA -1900

SEQIDNO10 - GGTCGTCTTGAGTCTCGTAAGAACAAACGTAACTTC-----ATCCTCGGA -1854

SEQIDNO7  - ATGGACAATACAACTGGAAGACGTACCATCCCCAAGTGCTTGTAACTGAT -1950

SEQIDNO10 - CGTGGTAGTTACGCCGGTGCCTATCGTTTTGCTGGTCTCTGGACTGGAGA -1904

SEQIDNO7  - ATGCGTTATGAGAATCATGGTCGGGAACCGATGGTCACTCAACGCAACAT -2000

SEQIDNO10 - TAACGCAAGTACGTGGGAATTCTGGAAGATTTCGGTCTCCCAAGTTCTTT -1954

SEQIDNO7  - TCATGCGTATACACTGTGCGAGTCTACTAGGAAGGAAGGGATCGTGGAAA -2050

SEQIDNO10 - CTCTAGGTCTCAATGGTGTGTGTATAGCGGGGTCTGATACGGGTGGTTTT -2004

SEQIDNO7  - ACGCAGACACTCTAACGAAGTTCCGCCGTAGCTACATTATCAGTCGTGGT -2100

SEQIDNO10 - GAGCCCGCACG-TACTGAGATTGGGGAGGAGAAATATTGCAGTCCGGAGC -2053

SEQIDNO7  - GGTTACATTGGTAACCAGCATTTCGGGGGTATGTGGGTGGGAGACAACTC -2150

SEQIDNO10 - TACTCATCAGGTGGTATACTGGATCATTCCTTTTGCCATGGCTTAGAAAC -2103

SEQIDNO7  - TACTACATCAAACTACATCCAAATGATGATTGCCAACAATATTAACATGA -2200

SEQIDNO10 - CACTACGTCAAGAAGGACAGGAAATGGTTCCAGGAACCATACGCGTACCC -2153

SEQIDNO7  - ATATGTCTTGCTTGCCTCTCGTCGGCTCCGACATTGGAGGATTCACCTCA -2250

SEQIDNO10 - CAAGCATCTTGAAACCCATCCAGAGCTCGCAGATCAAGCATGGCTTTACA -2203

SEQIDNO7  - TACGACAATGAGAA-TCAGCGAACGCCGTGTACCGGGGACTTGATGGTGA -2299

SEQIDNO10 - AATCTGTTCTAGAAATTTGCAGATACTGGGTAGAGCTAAGATATTCCCTC -2253

SEQIDNO7  - GGTATGTGCAGGCGGGCTGCCTGTTGCCGTGGTTCAGGAACCACTATGAT -2349

SEQIDNO10 - ATCCAGCTCCTTTACGACTGCATGTTCCAAAACGTGGTCGATGGTATGCC -2303

SEQIDNO7  - AGGTGGATCGAGTCCAAGGACCACGGAAAGGACTACCAGGAGCTGTACAT -2399

SEQIDNO10 - ACTTGCCAGATCTATGCTCTTGACCGATACTGAGGATACGACCTTCTTCA -2353

SEQIDNO7  - GTATCCGAATGAAATGGATACGTTGAGGAAGTTCGTTGAATTCCGTTATC -2449

SEQIDNO10 - ATGAGAGCCAAAAGTTCCTCGATAACCAATATATGGCTGGTGACGACATC -2403

SEQIDNO7  - GCTGGCAGGAAGTGTTGTACACGGCCATGTACCAGAATGCGGCTTTCGGA -2499
```

```
SEQIDNO10 - -CTTGTAGCACCCATCCTCCACAGCCGTAACGAGGTTCCGGGAGAGAACA -2452

SEQIDNO7  - AAGCCGATTATCAAGGCTGCTTCGATGTACAATAACGACTCAAACGTTCG -2549
SEQIDNO10 - GAGATGTCTATCTCCCTCTATTCCACACCTGGTACCC-CTCAAACTTGAG -2501

SEQIDNO7  - CAGGGCGCAGAACGATCATTTCCTTCTTGGTGGACATGATGGATATCGCA -2599
SEQIDNO10 - ACCGTGGGACGATCAGGGAGTCGCTTTAGGGAATCCTGTCGAAGGTGGCA -2551

SEQIDNO7  - TTCTGTGCGCGCCTGTTGTGTGGGAGAATTCGACCGAACGCGAATTGTAC -2649
SEQIDNO10 - GCGTTATCAACTACACTGCCAGGATTGTTGCCCCAGAGGATTATAATCTC -2601

SEQIDNO7  - TTGCCCGTGCTGACCCAATGGTACAAATTCGGTCCCGACTTTGACACCAA -2699
SEQIDNO10 - TTCCACAACGTGGTGCCGGTCTACATCAGAGAGGGTGCCATCATTCCGCA -2651

SEQIDNO7  - GCCTCTGGAAGGAGCGATGAACGGAGGGGACCGAATTTACAACTACCCTG -2749
SEQIDNO10 - AATTCAGGTACGCCAGTGGATTGGCGAAGGAGGGCCTAATCCCATCAAGT -2701

SEQIDNO7  - TACCGCAAAGTGAATCACCAATCTTCGTGAGAGAAGGTGCGATTCTCCCT -2799
SEQIDNO10 - TCAATATCTACCCTGGAAAGGACAAGGAGTATGTGACGTACCTTGATGAT -2751

SEQIDNO7  - ACCCGCTACACGTTGAACGGTGAAAACAAATCATTGAACACGTACACGGA -2849
SEQIDNO10 - GGTGTTAGCCGCGATAGTGCACCAGATGACCTCCCGCAGTACCGCGAGGC -2801

SEQIDNO7  - CGAAGATCCGTTGGTGTTTGAAGTATTCCCCCTCGGAAACAACCGTGCCG -2899
SEQIDNO10 - CTATGAGCAAGCGAAGGTCGAAGGCAAAGACGTCCAGAAGCAACTTGCGG -2851

SEQIDNO7  - ACGGTATGTGTTATCTTGATGATGGCGGTGTGACCACCAATGCTGAAGAC -2949
SEQIDNO10 - TCATTCAAGGGAATAAGACTAATGACTTCTCCGCCTCCGGGATTGATAAG -2901

SEQIDNO7  - AATGGCAAGTTCTCTGTCGTCAAGGTGGCAGCGGAGCAGGATGGTGGTAC -2999
SEQIDNO10 - GAGGCAAAGGGTTATCACCGCAAAGTTTCTATCAAACAGGAGTCAAAAGA -2951

SEQIDNO7  - GGAGACGATAACGTTTACGAATGATTGCTATGAGTACGTTTTCGGTGGAC -3049
SEQIDNO10 - CAAGACCCGTACTGTCACCATTGAGCCAAAACACAACGGATACGACCCCT -3001

SEQIDNO7  - CGTTCTACGTTCGAGTGCGCGGCGCTCAGTCGCCGTCGAACATCCACGTG -3099
SEQIDNO10 - CTAAGGAAGTTGGTAATTATTATACCATCATTCTTTGGTACGCACCGGGC -3051

SEQIDNO7  - TCTTCTGGAGCGGGTTCTCAGGACATGAAGGTGAGCTCTGCCACTTCCAG -3149
SEQIDNO10 - TTTGACGGCAGCATCGTCGATGTGAGCCAGGCGACCGTGAACATCGAGGG -3101
```

```
SEQIDNO7   - GGCTGCGCTGTTCAATGACGGGGAGAACGGTGATTTCTGGGTTGACCAGG -3199
              |  | |   | | |        |||||   |  ||    |   ||
SEQIDNO10 - CGGGGTGGAATGCGAAATTTTCAAGAACACCGGCTT----GCATACGGTT -3147
```

```
SEQIDNO7   - AGACAGATTCTCTGTGGCTGAAGTTGCCCAACGTTGTTCTCCCGGACGCT -3249
               |   |   |  .  |     |   |  |  ||   || |  |  ||
SEQIDNO10 - GTAGTCAACGTGAAAGAGGTGATCGGTACCACAAAGTCCGTCAAGATCAC -3197

SEQIDNO7   - GTGATCACAATTACCTAA -3267
              ||  ||| |   | ||
SEQIDNO10 - TTG--CACTACCGCTTAG -3213
```

Identity : 1119 (34.83%)
Number of gaps inserted in SEQIDNO7: 1
Number of gaps inserted in SEQIDNO10: 21

===03-MAY-2001===================NALIGN=========================PC/GENE===

===03-MAY-2001==================NALIGN========================PC/GENE===

```
*****************************************
* ALIGNMENT OF TWO NUCLEOTIDE SEQUENCES. *
*****************************************

The two sequences to be aligned are:

SEQIDNO7.
Total number of bases: 3267.

SEQIDNO11.
Total number of bases: 3279.

Open gap cost       : 50
Unit gap cost       : 20

The character to show that two aligned residues are identical is '|'

SEQIDNO7   - ATGTTTTCAACCCTTGCGTTTGTCGCACCTAGTGCGCTGGGAGCCAGTAC -50
             |||||| | |||||  | ||  | || || || |||||||  ||||| ||
SEQIDNO11  - ATGTTTCCTACCCTGACCTTCATAGCGCCCAGCGCGCTGGCCGCCAGCAC -50

SEQIDNO7   - CTTCGTAGGGGCGGAGGTCAGGTCAAATGTTCGTATCCATTCCGCTTTTC -100
             ||| || || |||||  || | ||     |||| || || |||||| |||
SEQIDNO11  - CTTTGTGGGCGCGGATATCCGATCGGGCATTCGCATTCAATCCGCTCTTC -100

SEQIDNO7   - CAGCTGTGCACACAGCTACTCGCAAAACCAATCGCCTCAATGTATCCATG -150
             | || |||| ||  |||   ||||  | ||| |    |||||||||||||
SEQIDNO11  - CGGCCGTGCGCAACGCTGTGCGCAGGAGCAAACATTACAATGTATCCATG -150

SEQIDNO7   - ACCGCATTGTCCGACAAACAAACGGCTACTGCGGGTAGTACAGACAATCC -200
             ||||||||||| ||||| ||||| ||||       | |  | ||||||||
SEQIDNO11  - ACCGCATTGTCTGACAAGCAAACCGCTATCAGTATTGGCCCTGACAATCC -200

SEQIDNO7   - GGACGGTATCGACTACAAGACCTACGATTACGTCGGAGTATGGGGTTTCA -250
             |||||||||| ||||| | | |||||||||| ||   |||  ||| || |
SEQIDNO11  - GGACGGTATCAACTACCAAAACTACGATTACATCCCTGTAGCGGGCTTTA -250

SEQIDNO7   - GCCCCCTCTCCAACACGAACTGGTTTGCTGCCGGCTCTTCTACCCCGGGT -300
               |||||||||||||| ||||||| ||||||||||||||| || |||||
SEQIDNO11  - CGCCCCTCTCCAACACCAACTGGTATGCTGCCGGCTCTTCCACTCCGGGC -300

SEQIDNO7   - GGCATCACTGATTGGACGGCTACAATGAATGTCAACTTCGACCGTATCGA -350
             |||||||| || ||||| ||||| ||||||||||| |||||||| || ||
SEQIDNO11  - GGCATCACCGACTGGACCGCTACCATGAATGTCAAATTCGACCGCATTGA -350

SEQIDNO7   - CAATCCGTCCATCACTGTCCAGCATCCCGTTCAGGTTCAGGTCACGTCAT -400
             |||||||||   | |     | ||||| |||||| ||||||||||||| |
SEQIDNO11  - CAATCCGTCGTACTCCAATAACCATCCTGTTCAGATTCAGGTCACGTCGT -400

SEQIDNO7   - ACAACAACAACAGCTACAGGGTTCGCTTCAACCCTGATGGCCCTATTCGT -450
             ||||||||||||||| |||| |||||||||||||||||||||| ||||||
SEQIDNO11  - ACAACAACAACAGCTTCAGGATTCGCTTCAACCCTGATGGCCCCATTCGT -450
```

```
SEQIDNO7   - GATGTGACTCGTGGGCCTATCCTCAAGCAGCAACTAGATTGGATTCGAAC -500
             || ||  |||| || |||||||| || ||||||||   |||||||||||
SEQIDNO11  - GACGTCTCTCGAGGACCTATCCTGAAACAGCAACTCACTTGGATTCGAAA -500

SEQIDNO7   - GCAGGAGCTGTCAGAGGGATGTGATCCCGGAATGACTTTCACATCAGAAG -550
              ||||||||| |  |||||||| ||||    ||||  ||| |  | ||||
SEQIDNO11  - CCAGGAGCTGGCGCAGGGATGTAATCCGAACATGAGCTTCTCTCCTGAAG -550

SEQIDNO7   - GTTTCTTGACTTTGAGACCAAGGATCTAAGCGTCATCATCTACGGAAAT -600
             |||| ||| ||||||| ||||| || |||| ||| || |||||||| ||
SEQIDNO11  - GTTTTTGTCTTTGAAACCAAAGACCTAAACGTTATAATCTACGGCAAC -600

SEQIDNO7   - TTCAAGACCAGAGTTACGAGAAAGTCTGACGGCAAGGTCATCATGGAAAA -650
             | |||||  ||||| ||||  |||    || ||| |  || ||||||| ||
SEQIDNO11  - TGCAAGATGAGAGTCACGAAGAAG---GATGGCTACCTCGTCATGGAGAA -647

SEQIDNO7   - TGATGAAGTTGGAACTGCATCGTCCGGGAACAAGTGCCGGGGATTGATGT -700
             ||| ||          |    |||     || || |||||  | ||||||||||
SEQIDNO11  - TGACGAGTGCAACTCGCAATCAGATGGCAATAAGTGTAGAGGATTGATGT -697

SEQIDNO7   - TCGTTGATAGATTATACGGTAACGCTATCGCTTCCGTCAACAAGAACTTC -750
              ||||||  |  |||||||||| ||||| ||||||||  | | ||| ||
SEQIDNO11  - ACGTTGACCGGCTATACGGTAATGCTATTGCTTCCGTACAAACGAATTTT -747

SEQIDNO7   - CGCAACGACGCGGTCAAGCAGGAGGGATTCTATGGTGCAGGTGAAGTCAA -800
             | ||| ||| |        | | |||  |||||||||||||||||||||||
SEQIDNO11  - CACAAAGACACTTCTCGGAACGAGAAATTCTATGGTGCAGGTGAAGTCAA -797

SEQIDNO7   - CTGTAAGTACCAGGAC---------------ACCTACATCTTAGAACGCA -835
             ||||   ||  ||||                ||  ||  |   ||||||||
SEQIDNO11  - CTGTCGCTATGAGGAGCAGGGTAAGGCGCCGACTTATGTTCTAGAACGCT -847

SEQIDNO7   - CTGGAATCGCCATGACAAATTACAACTACGATAACTTGAACTATAACCAG -885
             ||||| |||||||||| |||||||| ||||| ||||||||||| |||||
SEQIDNO11  - CTGGACTCGCCATGACCAATTACAATTACGACAACTTGAACTACAACCAA -897

SEQIDNO7   - TGGGACCTTAGACCTCCGCATCATGATGGTGCCCTCAACCCAGACTATTA -935
                ||| |    |||||   | ||       ||  | | ||  |||| ||
SEQIDNO11  - CCAGACGTCGTTCCTCCAGGTTATC------CCGACCATCCCAACTACTA -941

SEQIDNO7   - TATTCCAATGTACTACGCAGCACCTTGGTTGATCGTTAATGGATGCGCCG -985
              |||||||||||||||||||||| |||||| ||||| | |||||||| |
SEQIDNO11  - CATTCCAATGTACTACGCAGCACCGTGGTTGGTCGTTCAGGGATGCGCGG -991

SEQIDNO7   - GTACTTCGGAGCAGTACTCGTATGGATGGTTCATGGACAATGTCTCTCAA -1035
             | || ||| |||| |||||||| || ||||| ||||||||||||||||
SEQIDNO11  - GGACATCGAAGCAATACTCGTACGGTTGGTTTATGGACAATGTCTCTCAG -1041

SEQIDNO7   - TCTTACATGAATACTGGAGATACTACCTGGAATTCTGGACAAGAGGACCT -1085
             || |||||||| ||||||||||  | ||||| |  ||||| ||  ||||
SEQIDNO11  - TCGTACATGAACACTGGAGATACGGCGTGGAACTGCGGACAGGAAAACCT -1091

SEQIDNO7   - GGCATACATGGGCGCGCAGTATGGACCATTTGACCAACATTTTGTTTACG -1135
```

```
                 |||||||||||||||||| || || |||||||| || || ||||| || |
SEQIDNO11 - GGCATACATGGGCGCGCAATACGGGCCATTTGATCAGCACTTTGTGTATG -1141

SEQIDNO7  - GTGCTGGGGGTGGGATGGAATGTGTGGTCACAGCGTTCTCTCTTCTACAA -1185
            ||| ||| | |||  | |||  ||| |||| |||||||||  |||| |||
SEQIDNO11 - GTGATGGAGATGGCCTTGAAGATGTCGTCAAAGCGTTCTCCTTTCTTCAA -1191

SEQIDNO7  - GGCAAGGAGTTCGAGAACCAAGTTCTCAACAAACGTTCAGTAATGCCTCC -1235
            || |||||||||||  || ||   |||||||| ||||| |||||||||||
SEQIDNO11 - GGAAAGGAGTTCGAAGACAAAAAACTCAACAAGCGTTCTGTAATGCCTCC -1241

SEQIDNO7  - GAAATACGTCTTTGGTTTCTTCCAGGGTGTTTTCGGGACTTCTTCCTTGT -1285
            ||| ||||| ||||||||||||||||||||||||||  |   |||  |||
SEQIDNO11 - GAAGTACGTGTTTGGTTTCTTCCAGGGTGTTTTCGGTGCACTTTCACTGT -1291

SEQIDNO7  - TGAGAGCGCATATGCCAGCAGGTGAGAACAACATCTCAGTCGAAGAAATT -1335
            |||    | || |||| || || |||||||||||||||||  |||| |||
SEQIDNO11 - TGAAGCAGAATCTGCCTGCCGGAGAGAACAACATCTCAGTGCAAGAGATT -1341

SEQIDNO7  - GTAGAAGGTTATCAAAACAACAATTTCCCTTTCGAGGGGCTCGCTGTGGA -1385
            || || ||||| ||  | ||| | | ||| || || |||||||| || ||
SEQIDNO11 - GTGGAGGGTTACCAGGATAACGACTACCCCTTTGAAGGGCTCGCGGTAGA -1391

SEQIDNO7  - CGTGGATATGCAAGACAACTTGCGGGTGTTCACCACGAAGGGCGAATTTT -1435
             || |||||||||||  |  |||| ||||| || || ||    |||| ||
SEQIDNO11 - TGTTGATATGCAAGATGATCTGCGAGTGTTTACTACCAAACCAGAATATT -1441

SEQIDNO7  - GGACCGCAAACAGGGTGGGTACTGGCGGGGATCCAAACAACCGATCGGTT -1485
            || | ||||||| ||| ||    ||||| ||||| || ||| |||| || 
SEQIDNO11 - GGTCGGCAAACATGGTAGGCGAAGGCGGTGATCCTAATAACAGATCAGTC -1491

SEQIDNO7  - TTTGAATGGGCACATGACAAAGGCCTTGTTTGTCAGACAAATATAACTTG -1535
            |||||||||||||||||||  |||||||| |||||||| ||  |||||||
SEQIDNO11 - TTTGAATGGGCACATGACAGGGGCCTTGTCTGTCAGACGAACGTAACTTG -1541

SEQIDNO7  - CTTCCTGAGGAATGATAACGAGGGGCAAGACTACGAGGTCAATCAGACGT -1585
            |||| ||||||| ||||||   ||| ||   ||||| || |||||||| |
SEQIDNO11 - CTTCTTGAGGAACGATAACAGTGGGAAACCATACGAAGTGAATCAGACAT -1591

SEQIDNO7  - TAAGGGAGAGGCAGTTGTACACGAAGAACGACTCCCTGACGGGTACGGAT -1635
            | |||||||  |||||||| |||||||| || ||| |||     || |||
SEQIDNO11 - TGAGGGAGAAACAGTTGTATACGAAGAATGATTCCTTGAACAACACCGAT -1641

SEQIDNO7  - TTTGGAATGACCGACGACGGCCCCAGCGATGCGTACATCGGTCATCTGGA -1685
            |||||||  |||   || || ||  ||||||||||||| || ||| ||||
SEQIDNO11 - TTTGGAACTACCTCGGATGGGCCTGGCGATGCGTACATTGGACATTTGGA -1691

SEQIDNO7  - CTATGGGGGTGGAGTAGAATGTGATGCACTTTTCCCAGACTGGGGACGGC -1735
            |||||| |||||||| || ||||||||| | |||||||||||||| || |
SEQIDNO11 - CTATGGTGGTGGAGTGGAGTGTGATGCAATCTTCCCAGACTGGGGTCGAC -1741

SEQIDNO7  - CTGACGTGGCCGAATGGTGGGGAAATAACTATAAGAAACTGTTCAGCATT -1785
            | ||||||||  ||||||||||| | ||||| ||||| ||||||||||||
SEQIDNO11 - CAGACGTGGCTCAATGGTGGGGAGAAAACTACAAGAAGCTGTTCAGCATT -1791
```

```
SEQIDNO7  - GGTCTCGACTTCGTCTGGCAAGACATGACTGTTCCAGCAATGATGCCGCA -1835
            |||||||| ||||| ||||| || ||||| || || || |||||||||||
SEQIDNO11 - GGTCTCGATTTCGTGTGGCAGGATATGACGGTACCTGCGATGATGCCGCA -1841

SEQIDNO7  - CAAAATTGGCGATGACATCAATGTGAAACCGGATGG---------GAATT -1876
            |  | | || ||||   ||||    ||  | | | |           |   |
SEQIDNO11 - CCGACTCGGTGATGCTGTCAACAAAAATTCCGGTAGTTCGGCGCCGGGCT -1891

SEQIDNO7  - GGCCGAATGCGGACGATCCGTCCAATGGACAATACAACTGGAAGACGTAC -1926
            |||||||| | ||||||| ||||| |||| ||||||||||||   | ||
SEQIDNO11 - GGCCGAATGAGAACGATCCATCCAACGGACGATACAACTGGAAATCTTAT -1941

SEQIDNO7  - CATCCCCAAGTGCTTGTAACTGATATGCGTTATGAGA---ATCATGGTCG -1973
            ||||| |||||||| || || || ||||| ||||       |  ||||  |
SEQIDNO11 - CATCCGCAAGTGCTCGTGACCGACATGCGCTATGGTGCAGAGTATGGAAG -1991

SEQIDNO7  - GGAACCGATGGTCACTCAACGCAACATTCATGCGTATACACTGTGCGAGT -2023
            ||||||||||||   |||||||||||||||| || || || || || || |
SEQIDNO11 - GGAACCGATGGTGTCTCAACGCAACATTCACGCCTACACTCTTTGTGAAT -2041

SEQIDNO7  - CTACTAGGAAGGAAGGGATCGTGGAAAACGCAGACACTCTAACGAAGTTC -2073
            |||| ||    ||| || || |||| ||||||||||| | | || ||||||
SEQIDNO11 - CTACCAGACGGGAGGGAATTGTGGGAAACGCAGACAGTTTGACCAAGTTC -2091

SEQIDNO7  - CGCCGTAGCTACATTATCAGTCGTGGTGGTTACATTGGTAACCAGCATTT -2123
            ||||| || ||||| |||||||| || |||||||| ||||||||||||||
SEQIDNO11 - CGCCGCAGTTACATCATCAGTCGAGGAGGTTACATCGGTAACCAGCATTT -2141

SEQIDNO7  - CGGGGGTATGTGGGTGGGAGACAACTCTACTACATCAAACTACATCCAAA -2173
            ||| || |||||||| || ||||||   | | |||   |   |||| ||||||
SEQIDNO11 - CGGAGGGATGTGGGTTGGGGACAACAGTGCCACAGAATCCTACCTCCAAA -2191

SEQIDNO7  - TGATGATTGCCAACAATATTAACATGAATATGTCTTGCTTGCCTCTCGTC -2223
            ||||| | || |||| ||| ||||||||||||| ||| | || || ||
SEQIDNO11 - TGATGTTGGCGAACATTATCAACATGAATATGTCGTGCCTCCCGCTAGTT -2241

SEQIDNO7  - GGCTCCGACATTGGAGGATTCACCTCATACGACAATGAGAATCAGCGAAC -2273
            ||||| || ||||| || ||||||    ||| |  ||| |    | | |||
SEQIDNO11 - GGCTCTGATATTGGCGGGTTCACCCAGTACAATGATGCGGGCGACCCAAC -2291

SEQIDNO7  - GCCGTGTACCGGGGACTTGATGGTGAGGTATGTGCAGGCGGGCTGCCTGT -2323
             ||       || ||| |||||||| || |   |||||||| ||||| |||
SEQIDNO11 - CCC------CGAGGATTTGATGGTAAGATTCGTGCAGGCTGGCTGTCTGC -2335

SEQIDNO7  - TGCCGTGGTTCAGGAACCACTATGATAGGTGGATCGAGTCCAAGGACCAC -2373
            | ||||||||||| ||||||||||| |||||||| ||||||||| | |||
SEQIDNO11 - TACCGTGGTTCAGAAACCACTATGACAGGTGGATTGAGTCCAAGAAGCAC -2385

SEQIDNO7  - GGAAAGGACTACCAGGAGCTGTACATGTATCCGAATGAAATGGATACGTT -2423
            || ||| | ||||||||| | |||||||| |||      ||| |||||||||
SEQIDNO11 - GGGAAGAAATACCAGGAGTTATACATGTACCCGGGGCAAAAGGATACGTT -2435

SEQIDNO7  - GAGGAAGTTCGTTGAATTCCGTTATCGCTGGCAGGAAGTGTTGTACACGG -2473
```

```
SEQIDNO11 - GAAGAAGTTCGTTGAATTCCGCTACCGCTGGCAGGAGGTTTTGTACACAG -2485

SEQIDNO7  - CCATGTACCAGAATGCGGCTTTCGGAAAGCCGATTATCAAGGCTGCTTCG -2523
SEQIDNO11 - CCATGTACCAAAATGCTACCACTGGAGAGCCGATCATCAAGGCGGCGCCC -2535

SEQIDNO7  - ATGTACAATAACGACTCAAACGTTCGCAGGGCGCAGAACGATCATTTCCT -2573
SEQIDNO11 - ATGTACAACAACGACGTCAACGTGTATAAATCGCAGAATGATCATTTCCT -2585

SEQIDNO7  - TCTTGGTGGACATGATGGATATCGCATTCTGTGCGCGCCTGTTGTGTGGG -2623
SEQIDNO11 - TCTCGGTGGACATGACGGCTATCGTATTCTCTGCGCACCTGTTGTGCGCG -2635

SEQIDNO7  - AGAATTCGACCGAACGCGAATTGTACTTGCCCGTGCTGACCCAATGGTAC -2673
SEQIDNO11 - AAAATGCGACAAGTCGCGAAGTGTACCTGCCTGTGTATAGCAAGTGGTTC -2685

SEQIDNO7  - AAATTCGGTCCCGACTTTGACACCAAGCCTCTGGAAGGAGCGATGAACGG -2723
SEQIDNO11 - AAATTCGGACCGGACTTTGACACTAAGCCCTTGGAAAATGAGATTCAAGG -2735

SEQIDNO7  - AGGGGACCGAATTTACAACTACCCTGTACCGCAAAGTGAATCACCAATCT -2773
SEQIDNO11 - AGGTCAGACGCTTTATAATTACGCTGCACCGCTGAACGATTCGCCGATAT -2785

SEQIDNO7  - TCGTGAGAGAAGGTGCGATTCTCCCTACCCGCTACACGTTGAACGGTGAA -2823
SEQIDNO11 - TTGTGAGGGAAGGGACTATTCTTCCGACACGGTACACGCTGGACGGTGTG -2835

SEQIDNO7  - AACAAATCATTGAACACGTACACGGACGAAGATCCGTTGGTGTTTGAAGT -2873
SEQIDNO11 - AACAAATCTATCAACACGTACACAGACAATGATCCGCTTGTATTTGAGCT -2885

SEQIDNO7  - ATTCCCCCTCGGAAACAACCGTGCCGACGGTATGTGTTATCTTGATGATG -2923
SEQIDNO11 - GTTCCCTCTCGAAAACAACCAGGCGCATGGCTTGTTCTATCATGATGATG -2935

SEQIDNO7  - GCGGTGTGACCACCAATGCTGAAGACAATGGCAAGTTCTCTGTCGTCAAG -2973
SEQIDNO11 - GCGGTGTCACCACCAACGCTGAAGACTTTGGCAAGTATTCTGTGATCAGT -2985

SEQIDNO7  - GTGGCAGCGGAGCAGGATGGTGGTACGGAGACGATAACGTTTACGAATGA -3023
SEQIDNO11 - GTGAAGGCCGCGCAGGAAGGTTCTCAAATGAGTGTCAAGTTTGACAATGA -3035

SEQIDNO7  - TTGCTATGAGTACGTTTTCGGTGGACCGTTCTACGTTCGAGTGCGCGGCG -3073
SEQIDNO11 - AGTTTATGAACACCAATGGGGAGCATCGTTCTATGTTCGTGTTCGTAATA -3085

SEQIDNO7  - CTCAGTCGCCGTCGAACATCCACGTGTCTTCTGGAGCGGGTTCTCAGGAC -3123
SEQIDNO11 - TGGGTGCTCCGTCTAACATCAACGTATCTTCTCAGATTGGTCAACAGGAC -3135
```

```
SEQIDNO7  - ATGAAGGTGAGCTCTGCCACTTCCAGGGCTGCGCTGTTCAATGACGGGGA -3173
            ||| |   |||||| |  | |||||||||     |||||| |   |   |
SEQIDNO11 - ATGCAACAGAGCTCCGTGAGTTCCAGGGCGCAAATGTTCACTAGTGCTAA -3185

SEQIDNO7  - GAACGGTGATTTCTGGGTTGACCAGGAGACAGATTCTCTGTGGCTGAAGT -3223
              | || || | |||||||||||||   ||  | ||  ||||||| ||||
SEQIDNO11 - CGATGGCGAGTACTGGGTTGACCAGAGCACGAACTCGTTGTGGCTCAAGT -3235

SEQIDNO7  - TGCCCAACGTTGTTCTCCCGGACGCTGTGATCACAATTACCTAA -3267
            ||||    |  ||| |||  ||||||| ||||||  ||   | |
SEQIDNO11 - TGCCTGGTGCAGTTATCCAAGACGCTGCGATCACTGTTCGTTGA -3279


Identity  : 2450 (74.99%)
Number of gaps inserted in SEQIDNO7: 3
Number of gaps inserted in SEQIDNO11: 3

===03-MAY-2001==================NALIGN=======================PC/GENE===
```

===03-MAY-2001===================NALIGN=========================PC/GENE===

```
******************************************
* ALIGNMENT OF TWO NUCLEOTIDE SEQUENCES. *
******************************************

The two sequences to be aligned are:

SEQIDNO7.
Total number of bases: 3267.

SEQIDNO12.
Total number of bases: 1712.

Open gap cost      : 50
Unit gap cost      : 20

The character to show that two aligned residues are identical is '|'

SEQIDNO7  - ATGTTTTCAACCCTTGCGTTTGTCGCACCTAGTGCGCTGGGAGCCAGTAC -50
            |||
SEQIDNO12 - ATG----------------------------------------------- -3

SEQIDNO7  - CTTCGTAGGGGCGGAGGTCAGGTCAAATGTTCGTATCCATTCCGCTTTTC -100

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - CAGCTGTGCACACAGCTACTCGCAAAACCAATCGCCTCAATGTATCCATG -150

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - ACCGCATTGTCCGACAAACAAACGGCTACTGCGGGTAGTACAGACAATCC -200

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - GGACGGTATCGACTACAAGACCTACGATTACGTCGGAGTATGGGGTTTCA -250

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - GCCCCCTCTCCAACACGAACTGGTTTGCTGCCGGCTCTTCTACCCCGGGT -300

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - GGCATCACTGATTGGACGGCTACAATGAATGTCAACTTCGACCGTATCGA -350

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - CAATCCGTCCATCACTGTCCAGCATCCCGTTCAGGTTCAGGTCACGTCAT -400

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - ACAACAACAACAGCTACAGGGTTCGCTTCAACCCTGATGGCCCTATTCGT -450

SEQIDNO12 - -------------------------------------------------- -3
```

```
SEQIDNO7  - GATGTGACTCGTGGGCCTATCCTCAAGCAGCAACTAGATTGGATTCGAAC -500

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - GCAGGAGCTGTCAGAGGGATGTGATCCCGGAATGACTTTCACATCAGAAG -550

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - GTTTCTTGACTTTTGAGACCAAGGATCTAAGCGTCATCATCTACGGAAAT -600

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - TTCAAGACCAGAGTTACGAGAAAGTCTGACGGCAAGGTCATCATGGAAAA -650

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - TGATGAAGTTGGAACTGCATCGTCCGGGAACAAGTGCCGGGGATTGATGT -700

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - TCGTTGATAGATTATACGGTAACGCTATCGCTTCCGTCAACAAGAACTTC -750

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - CGCAACGACGCGGTCAAGCAGGAGGGATTCTATGGTGCAGGTGAAGTCAA -800

SEQIDNO12 - -------------------------------------------------- -3

SEQIDNO7  - CTGTAAGTACCAGGACACCTACATCTTAGAACGCACTGGAATCGCCATGA -850
                                                             |
SEQIDNO12 - -------------------------------------------------A -4

SEQIDNO7  - CAAATTACAACTACGATAACTTGAACTATAACCAGTGGGACCTTAGACCT -900
            |||| || || || || || |||||||| || ||   |||||| |  ||
SEQIDNO12 - CAAACTATAATTATGACAATTTGAACTACAATCAACCGGACCTCATCCCA -54

SEQIDNO7  - CCGCATCATGATGGTGCCCTCAACCCAGACTATTATATTCCAATGTACTA -950
            ||    ||||||   |      | || ||||| |||||||| |||||||
SEQIDNO12 - CCTGGCCATGATTCAG------ATCCTGACTACTATATTCCGATGTACTT -98

SEQIDNO7  - CGCAGCACCTTGGTTGATCGTTAATGGATGCGCCGGTACTTCGGAGCAGT -1000
             || ||||| ||| ||||||   ||||||     || ||   || ||||
SEQIDNO12 - TGCGGCACCATGGGTGATCGCACATGGATATCGTGGCACCAGCGACCAGT -148

SEQIDNO7  - ACTCGTATGGATGGTTCATGGACAATGTCTCTCAATCTTACATGAATACT -1050
            |||| |||||||||||  |||||||||| || || || ||||  || |||
SEQIDNO12 - ACTCTTATGGATGGTTTTTGGACAATGTATCCCAGTCCTACACAAACACT -198

SEQIDNO7  - GGAGATACTACCTGGAATTCTGGACAAGAGGACCTGGCATACATGGGCGC -1100
            || |||  | | |||    |||| ||  ||||  |||| |||||||| ||
SEQIDNO12 - GGCGATGATGCATGG---GCTGGTCAGAAGGATTTGGCGTACATGGGGGC -245

SEQIDNO7  - GCAGTATGGACCATTTGACCAACATTTTGTTTACGGTGCTGGGGGTGGGA -1150
```

```
SEQIDNO12 - ACAATGTGGGCCTTTCGATCAACATTTTGTGTATGAGGCTGGAGATGGAC -295

SEQIDNO7  - TGGAATGTGTGGTCACAGCGTTCTCTCTTCTACAAGGCAAGGAGTTCGAG -1200
SEQIDNO12 - TTGAAGACGTTGTGACCGCATTCTCTTATTTGCAAGGCAAGGAATATGAG -345

SEQIDNO7  - AACCAAGTTCTCAACAAACGTTCAGTAATGCCTCCGAAATACGTCTTTGG -1250
SEQIDNO12 - AACCAGGGACTGAATATACGTTCTGCAATGCCTCCGAAGTACGTTTTCGG -395

SEQIDNO7  - TTTCTTCCAGGGTGTTTTCGGGACTTCTTCCTTGTTGAGAGCGCATATGC -1300
SEQIDNO12 - ATTTTTCCAAGGCGTATTCGGAGCCACATCGCTGCTAAGGGACAACTTAC -445

SEQIDNO7  - CAGCAGGTGAGAACAACATCTCAGTCGAAGAAATTGTAGAAGGTTATCAA -1350
SEQIDNO12 - CTGCCGGCGAGAACAACGTCTCTTTGGAAGAAATTGTTGAAGGATATCAA -495

SEQIDNO7  - AACAACAATTTCCCTTTCGAGGGGCTCGCTGTGGACGTGGATATGCAAGA -1400
SEQIDNO12 - AATCAGAACGTGCCATTTGAAGGTCTTGCTGTGGATGTTGATATGCAAGA -545

SEQIDNO7  - CAACTTGCGGGTGTTCACCACGAAGGGCGAATTTTGGACCGCAAACAGGG -1450
SEQIDNO12 - TGACTTGAGAGTGTTCACTACGAGACCAGCGTTTTGGACGGCAAACAAGG -595

SEQIDNO7  - TGGGTACTGGCGGGGATCCAAACAACCGATCGGTTTTTGAATGGGCACAT -1500
SEQIDNO12 - TGGGGGAAGGCGGTGATCCAAACAACAAGTCAGTGTTTGAGTGGGCACAT -645

SEQIDNO7  - GACAAAGGCCTTGTTTGTCAGACAAATATAACTTGCTTCCTGAGGAATGA -1550
SEQIDNO12 - GACAGGGGCCTTGTCTGCCAGACGAATGTAACTTGCTTCTTGAAGAACGA -695

SEQIDNO7  - TAACGAGGGGCAAGACTACGAGGTCAATCAGACGTTAAGGGAGAGGCAGT -1600
SEQIDNO12 - GAAAAA------TCCTTACGAAGTGAATCAGTCATTGAGGGAGAAGCAGT -739

SEQIDNO7  - TGTACACGAAGAACGACTCCCTGACGGGTACGGATTTTGGAATGACCGAC -1650
SEQIDNO12 - TGTATACGAAGAGTGATTCCTTGGACAACATTGATTTTGGAACTACTCCA -789

SEQIDNO7  - GACGGCCCCAGCGATGCGTACATCGGTCATCTGGACTATGGGGGTGGAGT -1700
SEQIDNO12 - GATGGGCCTAGCGATGCGTACATTGGACACTTAGACTACGGTGGTGGTGT -839

SEQIDNO7  - AGAATGTGATGCACTTTTCCCAGACTGGGGACGGCCTGACGTGGCCGAAT -1750
SEQIDNO12 - GGAGTGTGATGCACTATTCCCAGACTGGGGTCGACCAGACGTGGCTCAAT -889

SEQIDNO7  - GGTGGGGAAATAACTATAAGAAACTGTTCAGCATTGGTCTCGACTTCGTC -1800
SEQIDNO12 - GGTGGGGCGATAACTACAAGAAACTATTCAGCATTGGTCTCGATTTCGTC -939
```

```
SEQIDNO7  - TGGCAAGACATGACTGTTCCAGCAATGATGCCGCACAAAATTGGCGATGA -1850

SEQIDNO12 - TGGCAAGATATGACGGTACCTGCGATGATGCCGCACCGACTCGGTGACCC -989

SEQIDNO7  - CATCAATGTGAAACCGGATGGGAAT---------TGGCCGAATGCGGACG -1891

SEQIDNO12 - TGTCGGCACAAATTCCGGTGAGACGGCGCCGGGCTGGCCGAATGATAAGG -1039

SEQIDNO7  - ATCCGTCCAATGGACAATACAACTGGAAGACGTACCATCCCCAAGTGCTT -1941

SEQIDNO12 - ATCCATCCAACGGACGATACAATTGGAAGTCTTACCATCCGCAAGTGCTC -1089

SEQIDNO7  - GTAACTGATATGCGTTATGAGAATCATGGTCGGGAACCGATGGTCACTCA -1991

SEQIDNO12 - GTGACTGACATGAGGTATGACGATTACGGAAGAGATCCCATTGTTACGCA -1139

SEQIDNO7  - ACGCAACATTCATGCGTATACACTGTGCGAGTCTACTAGGAAGGAAGGGA -2041

SEQIDNO12 - ACGCAATCTCCATGCCTACACTCTTTGTGAGTCTACTAGGAGGGAAGGCA -1189

SEQIDNO7  - TCGTGGAAAACGCAGACACTCTAACGAAGTTCCGCCGTAGCTACATTATC -2091

SEQIDNO12 - TTGTTGGAAACGCAGATAGTCTGACGAAGTTCCGCCGCAGCTATATTATC -1239

SEQIDNO7  - AGTCGTGGTGGTTACATTGGTAACCAGCATTTCGGGGTATGTGGGTGGG -2141

SEQIDNO12 - AGTCGTGGAGGCTACATCGGTAATCAGCACTTTGGTGGGATGTGGGTAGG -1289

SEQIDNO7  - AGACAACTCTACTACATCAAACTACATCCAAATGATGATTGCCAACAATA -2191

SEQIDNO12 - AGACAACTCTTCTACGGAAGACTACCTCGCAATGATGGTTATCAACGTTA -1339

SEQIDNO7  - TTAACATGAATATGTCTTGCTTGCCTCTCGTCGGCTCCGACATTGGAGGA -2241

SEQIDNO12 - TCAACATGAACATGTCCGGTGTCCCGCTCGTTGGTTCCGATATTGGAGGT -1389

SEQIDNO7  - TTCACCTCATACGACAATGAGAATCAGCGAACGCCGTGTACCGGGGACTT -2291

SEQIDNO12 - TTCACGGAGCATGACAA---------GAGAAACCCTTGCACACCGGACTT -1430

SEQIDNO7  - GATGGTGAGGTATGTGCAGGCGGGCTGCCTGTTGCCGTGGTTCAGGAACC -2341

SEQIDNO12 - GATGATGAGATTTGTGCAGGCTGGATGCTTGCTACCGTGGTTCAGGAACC -1480

SEQIDNO7  - ACTATGATAGGTGGATCGAGTCCAAGGACCACGGAAAGGACTACCAGGAG -2391

SEQIDNO12 - ACTACGATAGGTGGATCGAGAGCAAGAAACACGGAAAGAACTACCAAGAG -1530

SEQIDNO7  - CTGTACATGTATCCGAATGAAATGGATACGTTGAGGAAGTTCGTTGAATT -2441

SEQIDNO12 - TTGTACATGTACCGCGACCACTTGGACGCCTTGAGAAGTTTTGTGGAACT -1580

SEQIDNO7  - CCGTTATCGCTGGCAGGAAGTGTTGTACACGGCCATGTACCAGAATGCGG -2491
```

```
               ||| |||||||||||||||||||| ||||| |||||||| ||||||||
SEQIDNO12 - CCGCTATCGCTGGCAGGAAGTGTTATACACAGCCATGTATCAGAATGCTT -1630

SEQIDNO7  - CTTTCGGAAAGCCGATTATCAAGGCTGCTTCGATGTACAATAACGACTCA -2541
                ||| |||||||| |||||   | |  || |||||||| |||||
SEQIDNO12 - TGAACGGGAAGCCGATCATCAAAACGGTCTCCATGTACAACAACGATATG -1680

SEQIDNO7  - AACGTTCGCAGGGCGCAGAACGATCATTTCCTTCTTGGTGGACATGATGG -2591
            |||||         || ||||| || || |||||
SEQIDNO12 - AACGTCAAAGATGCTCAGAATGACCACTTCCT                   -1712

SEQIDNO7  - ATATCGCATTCTGTGCGCGCCTGTTGTGTGGGAGAATTCGACCGAACGCG -2641

SEQIDNO7  - AATTGTACTTGCCCGTGCTGACCCAATGGTACAAATTCGGTCCCGACTTT -2691

SEQIDNO7  - GACACCAAGCCTCTGGAAGGAGCGATGAACGGAGGGGACCGAATTTACAA -2741

SEQIDNO7  - CTACCCTGTACCGCAAAGTGAATCACCAATCTTCGTGAGAGAAGGTGCGA -2791

SEQIDNO7  - TTCTCCCTACCCGCTACACGTTGAACGGTGAAAACAAATCATTGAACACG -2841

SEQIDNO7  - TACACGGACGAAGATCCGTTGGTGTTTGAAGTATTCCCCCTCGGAAACAA -2891

SEQIDNO7  - CCGTGCCGACGGTATGTGTTATCTTGATGATGGCGGTGTGACCACCAATG -2941

SEQIDNO7  - CTGAAGACAATGGCAAGTTCTCTGTCGTCAAGGTGGCAGCGGAGCAGGAT -2991

SEQIDNO7  - GGTGGTACGGAGACGATAACGTTTACGAATGATTGCTATGAGTACGTTTT -3041

SEQIDNO7  - CGGTGGACCGTTCTACGTTCGAGTGCGCGGCGCTCAGTCGCCGTCGAACA -3091

SEQIDNO7  - TCCACGTGTCTTCTGGAGCGGGTTCTCAGGACATGAAGGTGAGCTCTGCC -3141

SEQIDNO7  - ACTTCCAGGGCTGCGCTGTTCAATGACGGGGAGAACGGTGATTTCTGGGT -3191

SEQIDNO7  - TGACCAGGAGACAGATTCTCTGTGGCTGAAGTTGCCCAACGTTGTTCTCC -3241

SEQIDNO7  - CGGACGCTGTGATCACAATTACCTAA -3267

Identity  : 1272 (74.30%)
Number of gaps inserted in SEQIDNO7: 1
Number of gaps inserted in SEQIDNO12: 5

===03-MAY-2001==================NALIGN=======================PC/GENE===
```